BIS

CONFERENCE PAPERS Vol. 4

MONETARY POLICY AND THE INFLATION PROCESS

July 1997

BANK FOR INTERNATIONAL SETTLEMENTS
Monetary and Economic Department
BASLE

TABLE OF CONTENTS

Introduction

The second meeting of Central Bank Model Builders was held at the BIS on 16th-17th January, 1997. Following the suggestions of participants in the first meeting, the broad agenda was:

Monetary policy and the inflation process

On the basis of suggestions of individual central banks and the papers submitted, three sub-themes emerged which were discussed in separate sessions. The first dealt with *Estimates of potential output, the NAIRU and inflation*, a theme which was addressed in five papers. The second session looked more closely on *The modelling of the wage and price formation process*, a sub-theme addressed by another five papers. The final session dealt with *Modelling monetary policy: responses, influences and effects* and contained a broader range of subjects, including, among others, the impact of inflation on growth, lags in the transmission from changes in monetary policy to output, the modelling of credibility effects and estimating money demand equations based on European-wide measures of monetary aggregates. The contributions (including discussants' comments) are reproduced in the order in which they were presented; the remainder of this introduction provides a summary of each paper, with the last section attempting to derive some broad conclusions.

1st Session: Estimates of potential output, the NAIRU and inflation

The paper by *P. St-Amant and S. van Norden (Bank of Canada)* reviews various methodologies for estimating potential output and the output gap that have recently been studied at the Bank of Canada, including applications to Canadian data.

The authors first discuss the Hodrick-Prescott (HP) filter which is often used to measure the output gap on the grounds that it extracts business cycle frequencies and can estimate an unobserved cyclical component. It is noted, however, that the filter is unlikely to do well for series such as real output, whose spectrum has the typical Granger shape; in addition it will often fail to adequately measure cyclical components. These problems are accentuated at the end of samples which is the period most relevant to policy makers. Finally, univariate filters will only provide information about the current output gap if the gap is Granger-caused by output growth which is not the case when potential output is exogenous.

Various extensions to the HP filter have incorporated additional information derived from economic relationships. These "hybrid" methods are driven by a desire to estimate output gaps that are not only conditioned by structural information but are "smooth" as well. However, they are hard to estimate, may not be robust to alternative calibrations and do not allow easy calculations of confidence intervals. Furthermore, one of the methods tested at the Bank does not perform as well as the simple HP filter in terms of isolating fluctuations of output originating from business cycle frequencies. Another approach discussed in the paper replaces the smoothness assumption with an unrestricted but linear filter.

The paper finally turns to multivariate filtering methods based on VARs incorporating long-run restrictions. Unlike univariate filters, such methods do not suffer from end-of-sample problems and can project values for the output gap. One advantage of the VAR-method using long-run restrictions over other multivariate methods, such as that of Beveridge and Nelson, is that it does not restrict the dynamics of potential output a priori. The paper investigates the implications of long-run restrictions for real output alone and for real output as well as inflation, arguing that the latter should be of interest to policymakers who focus on movements in real output associated with movements of trend inflation. Unfortunately, the applications considered involve rather wide confidence intervals, though using VARMAs instead of VARs may reduce that uncertainty.

In their paper on the link between the output gap and inflation in France, *J. Baude and G. Cette (Bank of France)* first discuss various reasons for analysing the cyclical position of an economy, including the identification of growth potential, inflationary pressures and the structural budget balance. The paper focuses on the second in deriving the relationship between the rate of inflation and the cyclical position of the economy, with the latter determined using various measures of the output gap that are all chosen to satisfy three criteria: easy to use, reproducible and rapid implementation.

The output gaps proposed range from measures based on single variables to a multivariate approach including a structural component. The results obtained are mutually consistent and also consistent with those of other studies. As a novel feature, the paper defines equilibrium unemployment as the rate at which firms' short-term profit shares are stable and calculates the equilibrium rate by relating wage pressures to assumed pressures in the labour marker, with the latter proxied by the difference between the actual and the equilibrium unemployment rate. This definition also enables the authors to identify the influence of changes in firms' net interest payments on their profit share and the NAIRU. Like other indicators in this area, those derived in this paper are conditional on a wide range of assumptions, some of which may be open to question. In its empirical part, the paper, shows that the output gap and the rate of capacity utilisation influence price inflation with broadly similar amplitudes. Moreover, the response pattern is stable and does not seem to depend on whether inflation is high or low.

P.G. Fisher, L. Mahadeva and J.D. Whitley (Bank of England) start by pointing out that, even though the concept of an output gap is relatively straightforward and its importance for understanding inflationary pressures is quite clear, it is difficult to measure accurately. Different measures of the output gap are illustrated for the United Kingdom and, in the face of this wide dispersion of estimates, the paper argues that it is preferable to adopt a procedure which has an economic theory content. Hence it uses a two-factor production function approach which also lends itself to deriving a statistical measure of uncertainty around the point estimates. The paper further argues that, while the output gap is extremely useful in describing the inflationary transmission mechanism, its practical relevance depends on how robust and well-defined is the empirical relationship between inflation and measures of the output gap.

Against this background, the authors derive an explicit measure of inflation expectations and, by incorporating this into the model, show that their preferred measure of the output gap based on quarterly data is a reasonably good indicator of inflation, in contrast to measures based on annual data. It is further shown that there is little evidence of changes in labour market flexibility but some signs of a mild asymmetry and of sectoral bottlenecks. The main finding, however, is one of a reasonably robust relationship between the output gap and inflation which, after allowing for supply-side shocks, does not depend critically on particular point estimates of the output gap. In this way, the output gap can provide a useful adjunct to other approaches used in the Bank's forecasting process. However, a complete evaluation needs to allow for the endogeneity of the output gap concept itself since it can be considered part of a wider system and not as a reduced form of the whole system.

In his paper on output gap and inflation in Japan, *T. Watanabe (Bank of Japan)* first estimates potential GDP, for which a Cobb-Douglas function with constant returns to scale is used. For total factor productivity an upward shift for the "bubble" period is identified but the 1990s pose a problem because of the end-point problem discussed in the paper by St-Amant and van Norden; in the end, Watanabe imposes a return to pre-bubble productivity growth. Potential input of labour is calculated using trend changes in participation rates for various age groups, the maximum number of overtime hours and a trend decline in the length of the normal work week. While the paper notes that "structural changes" are likely to have occurred in the labour market, these are not taken into account. Potential capital input is calculated taking the total capital stock for manufacturing times the maximum utilisation rate. Capital stock data for other sectors are not available and a constant utilisation rate has to be assumed, though the author recognises that this assumption could be a source of error. When this measure of potential output is compared with actual output, it appears that the output gap started to increase in the early 1990s and peaked at about 7% in 1995. Moreover, when changes in the gap are disaggregated into a final demand and an import component, the latter dominates recent changes due to the marked rise in import penetration of products from East Asian countries.

The paper then turns to the relationship between various measures of output slack and the rate of inflation and first looks at which of three possible Phillips-curve models best describes developments in Japan: a downward-sloping curve with a stable trade-off; a natural-rate or NAIRU model; or a relationship whereby *changes* in the gap is the principal determinant of inflation. Graphical analysis as well as econometric tests clearly support the first model. It very much appears as if the Phillips curve is downward-sloping for Japan and has been stable since the early 1980s. The last section of the paper discusses the costs of disinflation or, more particularly, whether the monetary authorities should aim for moderate or zero inflation. By comparing the 1990s with the 1980s, the paper concludes that the fall in the rate of inflation has been accompanied by a marked rise in real wages relative to productivity which is suggestive of a high degree of nominal rigidity. Although the author fails to find strong evidence that the sacrifice ratio increases as inflation approaches zero, it is recognised that, in view of the mixed evidence, these impressions are very preliminary and that the issue needs further research.

F. Fritzer and H. Glück (Austrian National Bank) apply a system approach to estimating the NAIRU, the output gap and wage and price equations for Austria. More specifically, they use a model, previously estimated by Coe and Krueger on German data, with three main features: first, because wages are determined in levels by the target real wage hypothesis, deviations of actual from equilibrium unemployment do not cause a continuous acceleration or deceleration of nominal wage changes; second, because of a trend rise in unemployment since the early 1980s, the unemployment rate equation is estimated in first differences; and third, output prices seem to dominate consumer prices in determining wages.

The five equations of the model are first estimated as single equations using OLS and then as a system, including cross-equation parameter constraints. While some parameters do change when moving to a system approach, the essential features of the model remain intact. These include a high degree of real wage flexibility, a low influence of foreign prices on domestic prices compared with the openness of the Austrian economy, and a trend rise in equilibrium unemployment which is largely attributed to a decline in the number of apprentices.

2nd Session: The modelling of the wage and price formation process

Because the rate of inflation is a stationary I(0) process in Switzerland while the rate of unemployment is non-stationary (I(1)), the two variables are statistically nonconforming. This is one of the reasons why *F. Ettlin (National Bank of Switzerland)* specifies and estimates his price adjustment equation for Switzerland as a relationship where the rate of inflation is mainly tied to current and lagged changes in import prices, the difference in productivity growth between the mainly goods-producing tradable sector and the mainly service-producing nontradable sector[1] and *changes* in the rate of unemployment.

Initially, a simple version of this model is estimated for comparison with a corresponding Phillips-curve specification and the empirical results support the model chosen by the author. In particular, while changes in unemployment are significant, the level is not only insignificant but also has the wrong sign. Subsequently, the model is estimated by a two-stage cointegration and error correction procedure, which further strengthens the support of the model chosen. Finally, by

---

[1] The price equation is essentially derived from the Scandinavian model of inflation which is driven by price changes in the sectors exposed to international competition and by the typically large differential in productivity developments between the exposed and the sheltered sectors.

decomposing inflation in Switzerland since the mid-1970s, the paper documents that changes in import prices were the dominant cyclical influence and the second most important contributor to inflation. The largest part is attributed to the substantial difference in sectoral productivity growth. In contrast, the contribution of changes in unemployment is relatively modest.

In their model of the wage and price formation in Belgium, *M. Dombrecht and P. Moës (National Bank of Belgium)* rely on the Layard-Nickell approach but with the important addition that producers' demand is driven by an "Almost Ideal Demand System". This extension implies a non-constant own price elasticity of demand and a mark-up which depends on the real effective exchange rate. It further implies that the real exchange rate also affects the wage bargaining process, not because of myopia on the part of wage earners but via the link between the price elasticity of demand and the exchange rate.

Following the derivation of the theoretical model and the specifications, the empirical results are presented, including short-run wage and price equations as well as a steady-state reduced-form price equation. Among the findings in the paper, it is worth highlighting that even though the estimated equilibrium rate of unemployment seems to have fluctuated around 8¼% since the late 1970s, the wage and price equations do not validate the NAIRU model. Indeed, because of the link to the real effective exchange rate, the equilibrium level of unemployment depends on the degree of competitiveness, a result that is also typical for several other smaller economies. The exchange rate link also means that foreign prices have had a dominating influence on inflation in Belgium while the most important domestic influences on steady-state inflation are the growth of the per capita capital stock and the slowdown in total factor productivity growth since the mid-1970s.

The principal aim of the paper by *S. Fabiani, A. Locarno and P. Sestito (Bank of Italy) and G. Oneto (ISCO)* is to investigate, by a number of econometric tests, whether the recent incomes policy measures in Italy have reduced the NAIRU as well as the costs of bringing down inflation to the level of other European countries. Given the comprehensive nature of the measures, a complete answer to this question would require a rather broad framework, including the effects of incomes policies on the future bargaining structure and on the implementation of fiscal consolidation. However, due to the rather short time horizon and several unsettled issues about the bargaining structure, the authors limit their analysis to the narrower questions of the NAIRU and the costs of disinflation.

Following a short summary of the history of wage indexation in Italy, the paper turns to the various factors leading up to the dismantling of wage indexation as well as the introduction of two new elements in the 1993 agreement: better coordination and a pivotal role for future inflation in shaping nominal wage changes. In theory, these measures are likely to reduce the short-run response of nominal wages to prices[2] and thus increase the degree of nominal rigidity in the system. In contrast, possible effects on the NAIRU and on the degree of relative wage flexibility (notably between different regions) are more uncertain.

Essentially, the paper presents three separate empirical tests. First, a traditional Phillips curve is estimated and tested for parameter stability and out-of-sample forecasting ability. There are no signs of overall parameter instability. Moreover, even though the estimates are consistent with a fall in the NAIRU, the wide confidence interval precludes precise judgements. In contrast, there is significant evidence of a slower response of nominal wages to prices; the one-step ahead forecast errors, while not too large on statistical grounds, show a systematic time pattern, with a pronounced slowing of wage dynamics immediately after the agreements followed by a significant catch-up later on.

The paper next presents estimates of a two-regime Phillips curve and these further support the hypothesis of a significant shift in the adjustment of nominal wages to prices even though

2 It is stressed, however, that the price response could merely be delayed to the next renegotiations.

some of the parameters of the post-agreement regime are higher than expected. The third test consists of ex post and counter-factual simulations on the Bank's quarterly model (summarised in an Annex to the paper) for the period 1993-97. The simulations are generated by using residuals from the wage equations (the time pattern of which is indicative of a temporary wage moderation episode) as shocks to baseline developments. Three different scenarios are presented (Table 4), with each characterised by the response of the central bank to the absence of wage moderation. All in all, the simulations suggest that wage moderation significantly helped disinflation by limiting the associated output and employment costs, notably in conditions when the fiscal situation did not allow any compensatory measures.

One feature of the current US expansion is the absence of any acceleration in inflation, despite rapid output growth and a rate of unemployment well below previous estimates of the NAIRU. Against this background, the paper by *C. S. Lown and R. W. Rich (Federal Reserve Bank of New York)* analyses whether the apparent "inflation puzzle" reflects a fundamental shift in the inflation process or whether forecasters have simply been wrong, thereby creating a perception of a puzzle which does not exist.

The paper first reviews the recent behaviour of prices, including factors that have helped to mute the rate of inflation, and then specifies and estimates a price-inflation Phillips curve. After subjecting the estimated equation to various stability tests, the authors conclude that there is little evidence of any fundamental shifts. The paper then turns to some of the underlying determinants to see if they have behaved abnormally. In this context a Phillips curve for compensation is estimated and this reveals some evidence of a temporary shortfall in the growth of unit labour costs, related to the behaviour of compensation over the period 1992-94. To understand the sources of this shortfall, Lown and Rich look at a number of additional variables as well as various hypotheses concerning possible changes in labour market behaviour. Only one series (job leavers as a percentage of the civilian labour force) seem to have played a role but, since its contribution to explaining the shortfall is rather modest, the slowdown in compensation growth remains a topic for further research.

Taking the history of disinflation in the United States and the uniquely low costs associated with the Volcker disinflation episode of the early 1980s as a starting point, *A. Bomfim, R. Tetlow, P. von zur Muehlen and J. Williams (Board of Governors of the Federal Reserve System)* more generally address the costs of disinflation, based on simulations of the Board's new quarterly macroeconometric model (FRB/US). More specifically, they attempt to find the sources of costly inflation adjustment and how they depend on the formation of expectations of inflation, the speed with which agents learn of a change in the inflation target set by the central bank, and the way in which monetary policy changes are implemented.

The authors begin by briefly describing the new model, focusing on those aspects and features that are pertinent to the issue at hand. One novel feature is the introduction of polynomial adjustment costs which disentangle intrinsic dynamics of the model from expectational dynamics due to errors in the formation of expectations and the process of learning. The paper then discusses wage and price dynamics and the formation of expectations which may be either model consistent or, in a limited information version, based on a simple VAR. Moreover, both may be combined with various speeds of agents' learning of changes in monetary policy targets (Table 5). Monetary policy is modelled by a reaction function for the Federal funds rate and, among the options available, the paper focuses on two rules, distinguished by their degree of aggressiveness, one of which is the well-known Taylor rule.

Among the many and interesting empirical findings reported in the paper, two seem particularly relevant. First, on the choice between a gradual and a "cold shower" approach to disinflation, the paper finds that the latter has the smallest costs when expectations are adjusted sluggishly whereas, in the case of model consistent expectations and fast learning, a gradual approach seems the least costly. However, given the myriad of learning rules that might be adopted, the authors conclude that it is still an open question which choice is optimal. Secondly, an attempt to calibrate the model so that it produces sacrifice ratios of around 2-2½ (which appears to be the consensus view) shows that only few combinations of expectation formation methods and rules satisfy this criterion. In fact, even though the aggressive post-1970 rule includes the Volcker episode, most combinations produce sacrifice ratios of 3 or higher. It is possible, however, that sacrifice ratios in the literature, are calculated for periods that are too short to capture all the costs of disinflation.

3rd Session: Modelling monetary policy: responses, influences and effects

This session started with the paper by *J. Andrés, R. Mestre and J. Vallés (Bank of Spain)* which simulates the effects of various changes in monetary policy using an aggregate macroeconomic model. The paper first reviews the various transmission channels of monetary policy, starting with a change in a monetary aggregate and then turning to changes in interest rates, highlighting three principal differences: a *permanent* reduction in monetary aggregate growth is equivalent to a *temporary* increase in interest rates; to emulate the effects of a reduction in monetary growth, the equivalent change in interest rates requires an initial increase, followed by a reduction later on; and contrary to an intermediate aggregate target, a framework based on interest rate control lacks a nominal anchor (further explored in Annex B of the paper).

The paper next presents the model (the corresponding empirical estimates given in Annex A) which is constructed so as to reflect the main features of the transmission mechanism, with monetary policy defined as a short-term interest rate path aimed at achieving the Bank's inflation target. Using the model, the authors simulate various strategies to achieve a permanent reduction in the rate of inflation to that of Spain's major trading partners. Judging by the results, the most credible and least costly way to lower inflation is by temporarily increasing interest rates, whereby an appreciation of the exchange rate reinforces the process of disinflation by reducing expectations of inflation. Indeed, even though all strategies produce identical long-run effects, the time path of inflation and output depends crucially on changes in competitiveness as the exchange rate emerges as a principal component of the transmission mechanism. However, because the response of the exchange rate to changes in the short-term interest rate had to be imposed, the authors also recognise that further analysis of this issue constitutes an important area of future research.

The length of the transmission from changes in monetary policy to output is a key issue to central banks and has been the subject of much research over the years. However, there are serious problems in isolating the lags with any precision. The paper by *D. Gruen, J. Romalis and N. Chandra (Reserve Bank of Australia)* analyses this problem in the case of Australia, using a simple model for output growth to estimate the lag and then examining possible sources of estimation biases.

The paper first describes the model which has, as one of its main features, a cointegrating relationship between Australian and US GDP. It then presents the estimates, which show that a 1 percentage point rise in the real short-term interest rate will reduce the growth of real GDP by about one-third of a point in both the first and the second year, followed by a minor reduction in the third year. However, a potentially important source of bias is the fact that, because monetary policy is forward-looking, OLS estimates of the GDP response to changes in the real interest rate will be understated. Moreover, most of this understatement is likely to be found in the first-year effect, because information relating to future years is more uncertain. The paper illustrates the size of the bias, using assumed correlations between current real interest rates and residuals of the estimated equation, and then introduces the novel idea of instrumenting the real interest rates, through a policy reaction function for two different regimes. Although the estimates tend to confirm the priors of the authors, the instrumental method suffers from the problem that the standard error rises compared with the OLS estimates.

The final section of the paper attempts to evaluate whether the lags of the transmission mechanism have remained stable over time and finds no evidence that the policy changes in the 1990s have led to any shortening of the lags. The estimation results are also compared with those of other models for Australia and for the US as well. Overall, it appears that all estimates from Australian models are "within the same ballpark" while those for the United States differ somewhat, depending on which model is used.

The paper by *W. Jahnke and B. Landau (Deutsche Bundesbank)* applies the Bank's multi-country model to assess the macroeconomic effects of the transition to a common monetary policy in EMU based on two alternative simulations: one where short-term interest rates converge towards German rates and a second one in which short-term interest rates converge towards an average European rate. Section I gives a brief review of the model while sections II and III present the principal results of the two alternative simulations. The main findings are that convergence towards the German interest rate will be accompanied by an expansionary effect on output in all countries, both directly and indirectly through exports, and inflation accelerates. In contrast, if interest rates converge towards an average European rate, output will expand in some countries (mainly those with high initial interest rates) but contract in others and inflation will fall, even in countries that already have a low rate. Although the precise response depends on how expectations are formed, it appears that structural differences and different initial conditions cause the transition to a common monetary policy in EMU to have divergent effects in the participating countries. In other words, national fiscal and structural policies have to ensure that monetary convergence is accompanied by convergence of output and prices.

The paper by *H. Dillén, T. Karlsson and J. Nilsson (Bank of Sweden)* uses a forward-looking macroeconomic model for Sweden to evaluate possible credibility effects during the period of disinflation and following Sweden's adoption of an inflation target. After a brief discussion of the nature of credibility, or lack of credibility, in the two policy regimes, the paper makes a distinction between exogenous credibility shocks and endogenous changes in credibility. This is further developed by relating credibility shocks respectively to expected depreciations and the long-term interest rate differential against Germany. It is noted that the risk premia in the latter case is 4-8 times higher than in the former case, a ratio which is confirmed later on.

Section 2 of the paper presents a clear and concise description of the model and its two sub-models: a steady-state component and a dynamic component, The speed of adjustment in the latter depends on structural factors (notably, slow adjustment in labour and product markets compared with financial markets) as well as on the nature of the expectation formation process and the policy reaction function. The paper then presents the simulation results, starting with the base case, followed by two alternatives characterised by different degrees of credibility. However, the authors note that the results may be subject to an upward bias because the simulations start from a long-run equilibrium situation which was not the case for Sweden in 1992. Subject to this caveat, a main conclusion of the paper is that credibility shocks affecting the nominal exchange rate cause substantial real effects, possibly as high as 2% of GDP. Moreover, such credibility shocks seem to go a long way in explaining the positive correlation between the exchange rate and interest rates in Sweden, notably during the early 1990s.

M. M. G. Fase and C. C. A. Winder (Netherlands Bank) address a key issue of monetary economics; viz. that of using econometric methods in estimating and testing money demand functions. However, they do so within a new institutional and constitutional framework by looking at money demand functions for the European Union as a whole. Moreover, one of the issues their paper addresses is how and why a sharp rise in broad money aggregates could coincide with falling inflation, a puzzle they solve by including net financial wealth in the money demand equation.

Following a brief description of the construction of the data (further explained in an Annex) and a preliminary graphical analysis of the relation between liquidity ratios and wealth/income ratios, the paper discusses the specification of the money demand equations and the restrictions imposed, with estimation results presented in Tables 1 and 2. There is clear evidence of a substantial impact of wealth on the demand for M2 and M3, whereas no influence on the demand for M1 is found. This result may explain the remarkable increase of the broad money aggregates over the last decade and also means that after taking the growth of wealth into account, recorded monetary expansion could be judged fairly modest. A similar conclusion follows from an analysis of Divisia aggregates, confirming that the strong growth of M2 and M3 should be attributed to portfolio investment considerations rather than to an expansionary monetary policy.

In a second paper from the *Bank of Spain, J. Andrés and I. Hernando* address a problem of principal interest to central bankers; viz. the long-run effects of inflation on growth. Although there is a general consensus among policy makers that inflation reduces long-run growth, empirical evidence in support of this view has not been very conclusive. Noting that one problem is the absence of a theoretical framework for analysing the effects of inflation, the authors choose the convergence equation as the basis for their empirical estimates. This approach allows for a variety of effects of inflation on growth, including those which reduce accumulation rates as well as those that undermine the efficiency with which productive factors are being used. It also allows inflation to affect both the level of per capita real income and its rate of growth and, given the fact that the direction of causality is not unambiguous, it lends itself naturally to testing the presumed inflation-growth link.

Following a brief discussion of the various costs of inflation and the theoretical model used, the paper presents the empirical estimates, based on 4-year averages for 24 OECD countries spanning the period 1960-95. The results suggest two channels by which inflation influences growth: a reduction in both the propensity to invest and the efficiency with which inputs are used. The estimates further indicate that the negative effect of inflation impinges upon the level of per capita income but not on the sustainable real rate of growth. The authors also test for non-linear inflation effects; using two alternative and equally plausible methods, they find that the elasticities of growth with respect to inflation are largely independent of the level of inflation. Finally, they address the problem that the inflation coefficients obtained in most growth equations are likely to be biased because a demand-induced rise in growth has a positive effect on inflation. However, extensive causality tests mostly confirm that the negative impact of inflation on growth is robust as long as the evaluation is confined to the *sign*, whereas the *size* of the coefficients differs across specifications and samples.

Overall, the authors find that, even though the estimated coefficients are not always significant and vary in size, the negative effect of inflation on growth "survives" in most specifications, including the inclusion of additional regressors and country-specific effects. They do not find a permanent growth rate effect. However, due to the long convergence period, their "ballpark" estimate that 1 percentage point more inflation reduces the long-run per capita real income level by about 2% implies that annual growth will be about 0.07% lower for an extended period.

Similar to the paper by Jahnke and Landau, *D. Bowman and J. H. Rogers (Board of Governors of the Federal Reserve System)* relies on simulations in analysing the macroeconomic effects of transition towards EMU. More specifically, they use the Federal Reserve Board's global model to investigate three major fiscal policy issues: achieving budget balance through either spending cuts or tax increases; a permanent reduction in the debt/GDP ratio; and imperfect credibility of announced fiscal measures. The paper also analyses the effects of two alternative monetary policy rules and a fiscal scenario motivated by pending future developments.

The main results of the simulations can be summarised in four points: first, fiscal consolidation through spending cuts is accompanied by larger macroeconomic effects than consolidation through tax increases; second, a permanent reduction in the debt/GDP ratio has a recessionary effect, depending on the size of the reduction and the speed with which it is implemented; third; the monetary policy response needed to offset the recessionary effects of fiscal consolidation is larger the less credible is the fiscal programme; and fourth, monetary policy in several countries is sufficient to completely offset the effects of fiscal measures implemented to satisfy deficit and debt criteria required for participation in EMU.

Concluding observations

About two-thirds of the papers presented at the meeting dealt with problems concerning the estimation and interpretation of output gaps or the NAIRU, which many central banks consider an important diagnostic tool in deciding whether the current state of the economy calls for changes in monetary policy. These concluding observations start with this issue and then turn to problems less directly related to this debate.

As a starting point, consider the following two equations:

(i) $dp_t = a\,(u_t - u^*) + b\,dp^e + c\,z_t + \varepsilon_t$, with $a \leq 0$ and $b \leq 1$

(ii) $dp_t = a'\,(y_t - y^*) + b'\,dp^e + c'\,z'_t + \varepsilon'_t$, with $a' \geq 0$ and $b' \leq 1$

where p denotes prices (in logs), u the rate of unemployment, u^* the NAIRU[3], y real output (in logs), y^* potential output (in logs), z and z' variables affecting inflation independently of the cycle, ε and ε' error terms, d the first difference operator and superscript 'e' expectations.

A first question is which of the above specifications is more consistent with the data, given the dynamic behaviour of dp, on the one hand, and $u - u^*$ and $y - y^*$, on the other. By construction, $y - y^*$ is stationary whereas for dp and $u - u^*$ the question is more difficult to answer. u^* is usually unknown and has to be determined from (i) and the behaviour of u.[4] Moreover, except for the United States, and even for sample periods of 30-40 years, only a few countries satisfy the condition that u returns to a stable equilibrium after being shocked, which implies that u is best approximated by a non-stationary I(1) process.[5] As regards dp, some recent studies have found evidence of weak mean reversion or "fractional integration" for samples covering the post-war period[6] and most empirical work assumes that dp can be approximated by an I(0) process. Consequently, the better specification would seem to require that dp be related to du and $(y - y^*)$ which is the approach adopted in the Swiss and Austrian papers. In the former, this choice is based on the time series properties of the variables and in the latter it is justified by a target real wage equation. Other papers (see below) do not estimate equation (i) but derive u^* from estimates of (ii) and find a trend rise in u^* since the early 1980s. This is, of course, observationally equivalent to finding a relationship between

---

3 Note that the NAIRU, proposed in Modigliani, F. and L. Papademos (19975): "Targets for monetary policy in the coming year", *Brookings Papers on Economic Activity*, pp. 141-63, is an empirical concept and may not be identical to the equilibrium or natural rate of unemployment as defined in Friedman, M. (1968): "The role of monetary Policy", *American Economic Review*, pp. 1-17 and Phelps, E. (1968): "Money wage dynamics and labour market equilibrium", *Journal of Political Economy*, pp. 678-711.

4 In some studies u^* is estimated independently and subsequently inserted into equations such (i). The finding in the Austrian paper that the rise in u^* can be attributed to a fall in the number of apprentices relative to the labour force can be interpreted in this light.

5 When estimating the equation $u_t = au_{t-1} + bu_{t-2}$ for the post-war period, $a+b$ is, for most non-US countries, insignificantly different from unity. In the *very* long run, u is, of course, stationary as it is limited to the range 0-1. However, for shorter periods, it is difficult to say whether the large autoregressive component reflects weak mean reversion (i.e. high degree of persistence) or complete hysteresis. Some have attempted to overcome this problem by distinguishing between a "shock free" NAIRU and a "shock-inclusive "NAIRU, with the former being far more stable than the latter. The estimates in Elmeskov, J. (1993): "High and persistent unemployment: assessment of the problem and its causes", *OECD Working Paper*, No. 132 may be considered examples of a shock-inclusive NAIRU.

6 For further discussion of this issue, see Gagnon, J. (1997): "Inflation regimes and inflation expectations", Board of Governors of the Federal Reserve System, *International Finance Discussion Paper*, No. 581.

dp and du (or to the hysteresis assumption of $u^* = u_{t-1}$). In contrast, changes in u^* do not seem to have been the source of the US inflation "puzzle" (New York Federal Reserve Bank).[7]

A second question is whether developments in z and z' cause only temporary disturbances or rather permanent shifts in the trade-off between inflation and "gaps". Several papers (Austria, Belgium, Switzerland, Japan and France) include import prices in explaining domestic price developments on the assumption that, in open economies, changes in foreign prices and in the exchange rate can have an important effect. In fact, in the case of Belgium, the NAIRU is not unique but depends on the real exchange rate. For most other countries, import prices are significant but with a coefficient well below the share of imports in domestic demand, implicitly suggesting that movements in foreign prices or in the exchange rate are either not fully passed through into domestic prices or that the pass-through process takes longer than the specification assumes.

Foreign price developments are not, however, the only source of permanent changes in z and z'. Fiscal policies, such as changes in "tax wedges", incomes policies and changes in the target for monetary policy, may also have permanent effects. While none of the papers explicitly discuss the influence of tax wedges, the possible effect of incomes policies on u^* and the degree of nominal rigidity (as measured by c and c' above) is at the core of the Italian paper. Similarly, the analysis of possible credibility effects, associated with the move to an inflation target in the Swedish paper and the simulations of a faster learning process (A. Bomfim et al., Federal Reserve Board), may be interpreted as attempts to determine whether such changes increase c and c' and thereby lower the costs of reducing inflation.[8]

A third question has to do with the nature of the long-run Phillips curve. While a consensus view among policy makers and most analysts is that the long-run Phillips curve is vertical, the existence of a long-run trade-off cannot be rejected in the case of Japan. In other countries, employees' real wage target (Austria) or the central bank's exchange rate target (Belgium) also implies that deviations of u from u^* do not lead to continuous decelerations or accelerations of inflation.

A fourth issue addressed by many papers concerns the proper measurement of u^* and y^* and thus the reliability of $u - u^*$ and $y - y^*$ as diagnostic tools. Given that u^* and y^* are not known, the above equations are usually estimated with intercept terms, denoted by e an e' respectively. If it is further assumed that c and $c' = 1$ and that z and z' involve only temporary shocks, u^* and y^*, defined as the levels of u and y respectively where $dp = dp^e$, can, in principle, be derived from equations (i) and (ii) as $u^* = -e/a$ and $y^* = e'/a'$. However, the preferred method in most papers was to estimate y^* separately and then either ignore u^* (United Kingdom and Canada) or link $u - u^*$ to $y - y^*$ via an Okun-equation (A. Bomfim et al.) or by some other method (France). A number of different methods for estimating y^* were discussed, notably in the Canadian and UK papers, but all of them were found to suffer from measurement problems of various kinds. Among these, the end-point problem seems to be the most serious because it affects precisely that period which is most important for the conduct of policies. At the same time, when faced with several more or less good indicators of the output gap, one way of selecting the most policy relevant measure is to incorporate it into equation (ii) above and tests its ability to explain and predict inflation (Canada and the United Kingdom).

---

7 However, even for the United States and assuming that the error term in (i) is normally distributed, estimates of u^* are not very precise. For instance, Staiger, D., J. Stock and M. Watson (1996): "How precise are estimates of the natural rate of unemployment?" *NBER Working Paper*, No. 5477 estimate u^* at 6.2 with a 95% confidence band of 5.1 to 7.7. Moreover, if the assumption of normality is dropped the band widens to 4.7-8.3 (see R. Chang (1997): "Is low unemployment inflationary?" Federal Reserve Bank of Atlanta, *Economic Review*, pp. 4-13. There is also some evidence of a gradual change in u^* for the United States; see Gordon, R. (1997): "The time-varying NAIRU and its implications for monetary policy", *Journal of Economic Perspectives*, pp. 11-32.

8 As explained in both the Italian and the Swedish papers, it may be more appropriate to analyse policy changes as regime shifts rather than as changes in z and z' or in c and c'. In the simulation results reported by A. Bomfim et al., c and c' may be broadly interpreted as reduced-form parameters of the wage-price bloc in the new FRB/US model.

A fifth question is implicit in the one discussed above. Can u^* and y^* move independently of each other and, if so, should $u - u^*$ or $y - y^*$ be the preferred indicator for policy makers. Except for the United States, where the Okun-equation appears to have been stable, most countries have experienced a marked rise in u^* independently of product market developments. As a consequence, several papers (Canada, France and the United Kingdom) indicated a preference for the output gap as the principal diagnostic tool for policy advice even though it was clearly recognised that the output gap was only one endogenous factor within a broader system.

A final problem is whether the relationship between inflation and either the output gap or the NAIRU is both log linear and symmetric as is assumed in (i) and (ii) above. The new Federal Reserve Board model is largely linear so that (except for sign) excess demand and excess supply have similar effects on the rate of inflation. It also appears that the Phillips curve is largely linear in Japan and that the costs of disinflation do not rise when inflation is already very low.[9] In contrast, the UK paper uncovers two sources of non-linearities. First, a positive output gap seems to increase inflation by more than a negative gap of the same size reduces inflation. This implies that actual output has to be held below potential if inflation is to be kept stable over time and that reducing inflation is less costly if done gradually. Secondly, in the United Kingdom excess demand pressures and associated inflationary risks appear to vary across sectors, with the services sector being particularly prone to inflation.

Despite the uncertainties and measurement issues discussed above, there will be occasions when policy actions are required to counter potential inflationary pressures or to reduce inflation to a more acceptable level. This raises another set of issues that were mostly discussed in the last session of the meeting: what are the principal transmission channels of changes in monetary policy; how quickly do such changes affect inflation and output; what are the costs (sacrifice ratios) in terms of lost output and employment; and do the gains from lower inflation exceed the costs associated with reducing inflation?

As noted above, foreign prices play an important role in the determination of domestic prices in small and open economies. Similarly, the response of the exchange rate to changes in monetary instruments is an important element in the transmission of monetary policy, not only with respect to the speed with which inflation is affected but also as regards the associated output costs (Spain and Sweden). In addition (A. Bomfim et al.), sacrifice ratios will be affected by the formation of expectations of inflation and the process by which agents learn of a change in monetary policy with respect to the inflation target. In this context, a change in the monetary policy regime can also influence the transmission process as a higher degree of credibility or a more rapid speed of learning will reduce the sacrifice ratio (Sweden and Bomfim et al.).[10]

---

9 In a subsequent working paper from the Bank of Japan, Kimura and Ueda find that nominal wages in Japan decline in response to a worsening of aggregate labour market conditions, even when inflation is very low. In contrast, faced with negative micro shocks, firms prefer to adjust overtime hours rather than to reduce nominal wages or the number of employees.

10 It is interesting to note that cross-country regressions relating sacrifice ratios to central bank independence and other factors usually find that the countries with the most independent central banks tend to have the highest sacrifice ratios. There are, however, several ways in which a move towards a higher degree of independence may affect sacrifice ratios. On the one hand, if such a change leads to a faster learning process because monetary policy is regarded as more credible, c and c' in the above equations will rise and this reduces the sacrifice ratio (A. Bomfim et al.). On the other hand, a higher degree of credibility may also reduce nominal contracting, facilitate signal extraction or reduce the frequency with which firms adjust prices and such influences will reduce a and a' and thus increase the sacrifice ratio. For further discussion, see Hutchison, M. and C. Walsh (1996): "Central bank institutional design and the output cost of disinflation: did the 1989 New Zealand Reserve Bank Act affect the inflation-output trade-off?" *Reserve Bank of New Zealand Discussion Paper*, No. 96/6.

Monetary policy affects inflation via changes in output as well but this channel is subject to longer lags than the transmission through exchange rate changes. Determining the length of the lags is, however, subject to several problems, including the fact that the estimated coefficient on the policy instrument will be biased if policies are forward looking and policy decisions are based on expected output changes (Australia).

Whatever the transmission channel and the lags, monetary policy changes aimed at reducing inflation, or pre-empting a future rise, will, in general, induce some costs. This raises the question whether a lower rate of inflation leads to a net rise in discounted real output. Even though the gains from lower inflation are usually assumed to be permanent, and the costs are only temporary if the long-run Phillips curve is vertical, the empirical evidence on this issue is not very conclusive. However, when an appropriate analytical framework is adopted, it does appear that the gains exceed the costs, as inflation seems to have a negative effect on real income levels. This is clear from the empirical evidence even when a number of additional factors are included and the dual causality between inflation and output is taken into account (Spain, second paper).

One important conclusion emerging from the many papers addressing the issue of estimating NAIRU and the output gap and their developments over time is the large influence attributed to institutional and country-specific factors. This is the case, not only in papers which compared the United States with Europe but also when comparisons were made between European countries. Emphasis on such differences was a feature of two of the papers that discussed various aspects of the transition towards EMU. For instance, the convergence towards a common rate of interest for monetary policy in Europe (Germany) could be accompanied by a divergence of real growth and inflation due to institutional and country-specific factors as well as different initial conditions. Similarly, fiscal measures adopted to meet the Maastricht criteria would lead to real divergences and, to the extent that these differences were to be neutralised by monetary policy, the changes required would differ across countries (Bowman and Rogers, Federal Reserve Board). In contrast, despite institutional and country-specific factors, and notwithstanding that money demand has become unstable in a large number of countries, it appears that money demand equations based on EU aggregates have remained largely stable (Netherlands). Moreover, once the significant rise in wealth is taken into account, the acceleration in the growth of broad monetary aggregates in recent years can be interpreted as a portfolio shift rather than a more expansionary monetary policy.

Participants in the meeting

Australia:	Mr. David W.R. GRUEN
Austria:	Mr. Friedrich FRITZER Mr. Heinz GLÜCK
Belgium:	Mr. Michel DOMBRECHT Mr. Philippe MOËS
Canada:	Mr. Pierre ST-AMANT Mr. Simon VAN NORDEN
France:	Mr. John BAUDE Mr. Gilbert CETTE
Germany:	Mr. Wilfried JAHNKE Ms. Bettina LANDAU
Italy:	Mr. Alberto LOCARNO Mr. Paolo SESTITO
Japan:	Mr. Tsutomu WATANABE
Netherlands:	Prof. Martin M.G. FASE Mr. Carlo C.A. WINDER
Spain:	Mr. Ignacio HERNANDO Mr. Javier VALLÉS
Sweden:	Mr. Hans DILLÉN Mr. Jonny NILSSON
Switzerland:	Mr. Franz ETTLIN Ms. Barbara LÜSCHER
United Kingdom:	Mr. Tony YATES Mr. Lavan MAHADEVA
United States:	Mr. Robert W. RICH *(New York)* Mr. John H. ROGERS *(Washington)* Mr. Robert J. TETLOW *(Washington)*
BIS:	Mr. Renato FILOSA (Chairman) Mr. Joseph BISIGNANO Mr. Zenta NAKAJIMA Mr. Palle ANDERSEN
Discussants:	Mr. Claudio Borio Mr. Ben Cohen Mr. Sean Craig Mr. Wilhelm Fritz Mr. Gabriele Galati Mr. Stefan Gerlach Mr. Steve Kamin Mr. Frank Smets Mr. Gregory Sutton Mr. Kostas Tsatsaronis

Measurement of the output gap: a discussion of recent research at the Bank of Canada

Pierre St-Amant and Simon van Norden[1]

Introduction

Most macroeconomic models that are used for forecasting and policy analysis require an estimate of potential output. For example, at the Bank of Canada, estimates of potential output are important inputs in different "Phillips curve" models and in the staff's Quarterly Projection Model, where the gap between actual and potential output is a key variable determining the evolution of prices and wages. A level of real GDP above potential (a positive output gap) will often be seen as a source of inflationary pressures and a signal that monetary authorities interested in avoiding an acceleration of inflation should tighten monetary conditions. A level of real GDP below potential (a negative output gap) will have the opposite implication.

The output gap can thus be defined as the component of real output that is associated with changes in inflation.[2] Note that gaps could be calculated in markets other than that for goods and services. For example, gaps in the labour market have frequently been calculated and authors such as Hendry (1995) present "money gaps."

Unfortunately, measuring the output gap is not an easy task. Different sets of assumptions can be used together with different econometric techniques to provide different measures of the output gap. One common assumption is that the output gap is some part of the transitory (cyclical) component of real output. The methods discussed in this paper make that assumption.

The first group of methods we consider are those which simply use some (implicit or explicit) assumptions about the dynamics of real output to identify the output gap. For example, if one believed that real output was composed of a stationary component and a simple log-linear trend, the output gap could be measured as the residuals of a regression of log output on a linear time trend. Unfortunately, such a simple model does not adequately describe the behaviour of output, and measuring the temporary component in more complex models is problematic.

In this paper, we will assume that real output is I(1); that is, that the level of output is subject to permanent shocks so there is no deterministic trend towards which output tends to revert.[3]

---

1 The authors wish to thank Chantal Dupasquier, Paul Fenton, Gabriele Galati, Alain Guay, Seamus Hogan, Irene Ip, Robert Lafrance, René Lalonde, David Longworth, Tiff Macklem, John Murray and Brian O'Reilly for useful comments and discussions. They also thank Jennifer Page and Rebecca Szetto for their excellent research assistance. Of course, since the authors are solely responsible for the paper's content, none of the aforementioned are responsible for any remaining errors. The views expressed in this study are those of the authors and do not necessarily represent those of the Bank of Canada.

2 To be more precise, we should take into account expected inflation, and therefore define the output gap with respect to changes in unexpected inflation. Some models also imply a relationship between the change in the gap and inflation.

3 This is the most common assumption in modern applied macroeconomics and is consistent with the view that real output can be permanently affected by shocks, such as technological innovations. An alternative view is that output is stationary around a time trend, but that this time trend is subject to occasional random changes in its slope and intercept. Evidence for such a view is discussed by Perron (1989) and Weber (1995). As detecting changes in the slope or intercept near the end of a sample is quite difficult, such models imply that one cannot reliably measure the current

Many approaches have been proposed to identify the permanent and cyclical components of real output in such models, such as those proposed by Hodrick and Prescott (1997), Watson (1986) or by Beveridge-Nelson (1981). The problem is that the measured cyclical component may differ considerably from method to method. Quah (1992) argues that this is an intrinsic problem and that "...without additional ad hoc restrictions those (univariate) characterizations are completely uninformative..."

These problems have not prevented the widespread use of Hodrick and Prescott's filter to identify the cyclical component of output.[4] Arguments commonly made to justify its use are that:

- it extracts the relevant business-cycle frequencies of output;
- it closely approximates the cyclical component implied by reasonable time-series models of output.

We examine these arguments in Sections 1.2 and 1.3. We also note that, unlike much of the literature on "detrending", the focus of the problem confronting policymakers is to estimate the deviation from trend at the end rather than the middle of a data sample.[5] We conclude that such methods are unlikely to be suitable for use in a policy context, and we discuss economic factors which limit our ability to estimate the current output gap.

An important class of alternatives to these univariate dynamic methods are those which combine their assumptions with information from assumed or "structural" relationships between the output gap and other economic variables, such as a Phillips Curve or Okun's Law. We examine some of these in Section 2. Among them are the multivariate HP filters (MHPF) proposed by Laxton and Tetlow (1992) and Butler (1996), which is the general approach currently used in the staff economic projection of the Canadian economy at the Bank of Canada. In Section 2, we note that calibration of the MHPF methods has been problematic and that despite the inclusion of structural information, their estimates of the output gap have wide confidence intervals. Spectral analysis of the Canadian output gap resulting from the application of the MHPH method also gives the "disturbing" result that it includes a very large proportion of cycles much longer than what is usually defined as being business cycles. A reaction to these methods is the "Trivial Optimal Filter that may be Useful" (TOFU) approach suggested by van Norden (1995), which replaces the HP smoothing problem with the simpler restriction of a constant linear filter. The TOFU approach has yet to be shown to be workable.

The third and final class of methods we consider uses multivariate rather than univariate dynamic relationships, often in combination with structural relationships from economic theory, to estimate output gaps as a particular transitory component of real output. Some of these are examined in Section 3. One example is the decomposition method suggested by Cochrane (CO, 1994). This method is based on the permanent income hypothesis and uses consumption to define the permanent component of output which can then be used as a measure of potential output. Multivariate extensions of the Beveridge-Nelson decomposition method (MBN) have also been proposed to identify the

---

deviations from trend. As we argue below that this is what policymakers wish to measure, adoption of the "breaking-trend" model per se is not a solution to the problems of measuring output gaps which we discuss below.

4 We will henceforth refer to this method as the HP filter, although Hodrick and Prescott note that their method is due to Whittaker (1923) and Henderson (1924). Also, although the Hodrick and Prescott article is to be published in 1997, their working paper dates from 1981.

5 This is an oversimplification. More accurately, policymakers will usually be most interested in expected *future* values of the output gap, particularly when these expectations are conditioned on specific policy actions. This is more demanding than simply estimating the output gap at the end of sample, so our discussion of the additional difficulties introduced by end-of-sample problems underestimates the true difficulty of the policy problem. For that reason, we think good end-of-sample performance is a necessary rather than a sufficient condition for reliable estimation of the deviation from trend.

permanent component of output (Evans and Reichlin, 1994). A major restriction, used by both the CO and the MBN methods, is that the permanent component of real output is a random walk.

Section 3.1 of this paper, which draws heavily from Dupasquier, Guay and St-Amant (1996), discusses the CO and MBN methodologies and compares them with a structural vector autoregression methodology based on long-run restrictions imposed on output (LRRO.) This method was proposed by Blanchard and Quah (1989), Shapiro and Watson (1988), and King et al. (1991). One characteristic of the LRRO approach is that it does not impose restrictions on the dynamics of the permanent component of output. Instead, it allows for a permanent component comprising an estimated diffusion process for permanent shocks that can differ from a random walk. The output gap then corresponds to the cyclical component of output excluding the diffusion process of permanent shocks which is instead assigned to potential output. Instead, it allows for a permanent component comprising an estimated diffusion process of permanent shocks which is instead assigned to potential output. Section 3.2 presents an application of the LRRO method to Canadian data.

In Section 3.3 (which draws from Lalonde, Page and St-Amant (forthcoming)) we present another methodology based on long-run restrictions imposed on a VAR that associates restrictions imposed to real output *and* inflation. The output gap is then a part of the cyclical component of real output that is consistent with changes in the trend of inflation.[6]

The final section concludes with some directions for future research.

1. The HP filter

In recent years, mechanical filters have frequently been used to identify permanent and cyclical components of time series. The most popular of these mechanical filters is that proposed by Hodrick and Prescott (1997). This section evaluates the basic Hodrick-Prescott (HP) filter's ability to provide a useful estimate of the output gap. Section 2 then discusses some extensions and alternatives to the basic HP filter that have recently been proposed.

Guay and St-Amant (1996) show that the HP filter does a poor job in terms of extracting business cycle frequencies from macroeconomic time series. As a consequence, it is not an adequate approach to estimating an output gap constrained to correspond to the business-cycle frequencies of real GDP. This is discussed in Section 1.2, where we further argue that constraining the output gap in that way is not very attractive in any case. Guay and St-Amant also show that the HP filter is likely to do a poor job in terms of extracting an output gap assumed to correspond to the unobserved cyclical component of real GDP. This is discussed in Section 1.3. In Section 1.4, we focus explicitly on the HP filter's end-of-sample problems and conclude that these raise further doubts about the appropriateness of using the HP filter to estimate the output gap. Finally, Section 1.5 investigates what economic theory has to say about the possible usefulness of filters for estimating output gaps at the end of sample.

Most of the arguments in this section of the paper are drawn from Guay and St-Amant (1996) and van Norden (1995). Note that Guay and St-Amant show that the main conclusions that they reach concerning the HP filter also apply to the band-pass filter proposed by Baxter and King (1995).

1.1 The optimization problem

The HP filter decomposes a time series y_t into additive components, a cyclical

---

[6] Lalonde, Page and St-Amant also present a method associating the output gap with changes in the trend of inflation but which does not impose that the output gap is stationary.

component y_t^c and a growth component y_t^g,

$$y_t = y_t^g + y_t^c \tag{1}$$

Applying the HP filter involves minimizing the variance of y_t^c subject to a penalty for the variation in the second difference of y_t^g. This is expressed in the following equation:

$$\left\{y_t^g\right\}_{t=0}^{T+1} = \arg\min \sum_{t=1}^{T}\left[\left(y_t - y_t^g\right)^2 + \lambda\left[\left(y_{t+1}^g - y_t^g\right) - \left(y_t^g - y_{t-1}^g\right)\right]^2\right] \tag{2}$$

where λ, the smoothness parameter, penalizes the variability in the growth component. The larger the value of λ, the smoother the growth component. As λ approaches infinity, the growth component corresponds to a linear time trend. For quarterly data, Hodrick and Prescott propose setting λ equal to 1,600. King and Rebelo (1993) show that the HP filter can render stationary any integrated process of up to the fourth order.

1.2 How well does the HP filter extract business cycle frequencies?

Authors such as Singleton (1988) have shown that the HP filter can provide an adequate approximation of a high-pass filter when it is applied to stationary time series. Here we need to introduce some elements of spectral analysis. A zero-mean stationary process has a Cramer representation such as:

$$y_t = \int_{-\pi}^{\pi} \varepsilon^{i\omega t} dz(\omega) \tag{3}$$

where $dz(\omega)$ is a complex value of orthogonal increments, i is the imaginary number $(-1)^{1/2}$ and ω is frequency measured in radians, i.e. $-\pi \le \omega \le \pi$ (see Priestley (1981), chapter 4). In turn, filtered time series can be expressed as:

$$y_t^f = \int_{-\pi}^{\pi} \alpha(\omega) e^{i\omega t} dz(\omega), \text{ with} \tag{4}$$

$$\alpha(\omega) = \sum_{h=-k}^{k} a_h e^{-i\omega h} \tag{5}$$

Equation (5) is the frequency response (Fourier transform) of the filter. That is, $\alpha(\omega)$ indicates the extent to which y_t^f responds to y_t at frequency ω and can be seen as the weight attached to the periodic component $e^{i\omega t}dz(\omega)$. In the case of symmetric filters, the Fourier transform is also called the gain of the filter.

An ideal high-pass filter would remove low-frequency, or long-cycle, components and allow high-frequency, or short-cycle, components to pass through so that $\alpha(\omega)=0$ for $|\omega| \le \omega^p$, where ω^p has some predetermined value and $\alpha(\omega)=1$ for $|\omega| > \omega^p$. Figure 1 shows the squared gain of the HP filter. Very high frequencies are left aside because we want to focus on business-cycle frequencies as defined by NBER researchers since Burns and Mitchell (1946); i.e. cycles lasting no less than 6 and no more than 32 quarters. We see that the squared gain is 0 at zero frequency and is close to 1 from around frequency $\pi/10$ (6 quarters) and up. On the basis of Figure 1, the HP filter would appear to be and adequate approximation of a high-pass filter in that it removes most low frequencies and passes through most higher frequencies including business-cycle frequencies.

Figure 1

Squared gain of the HP filter

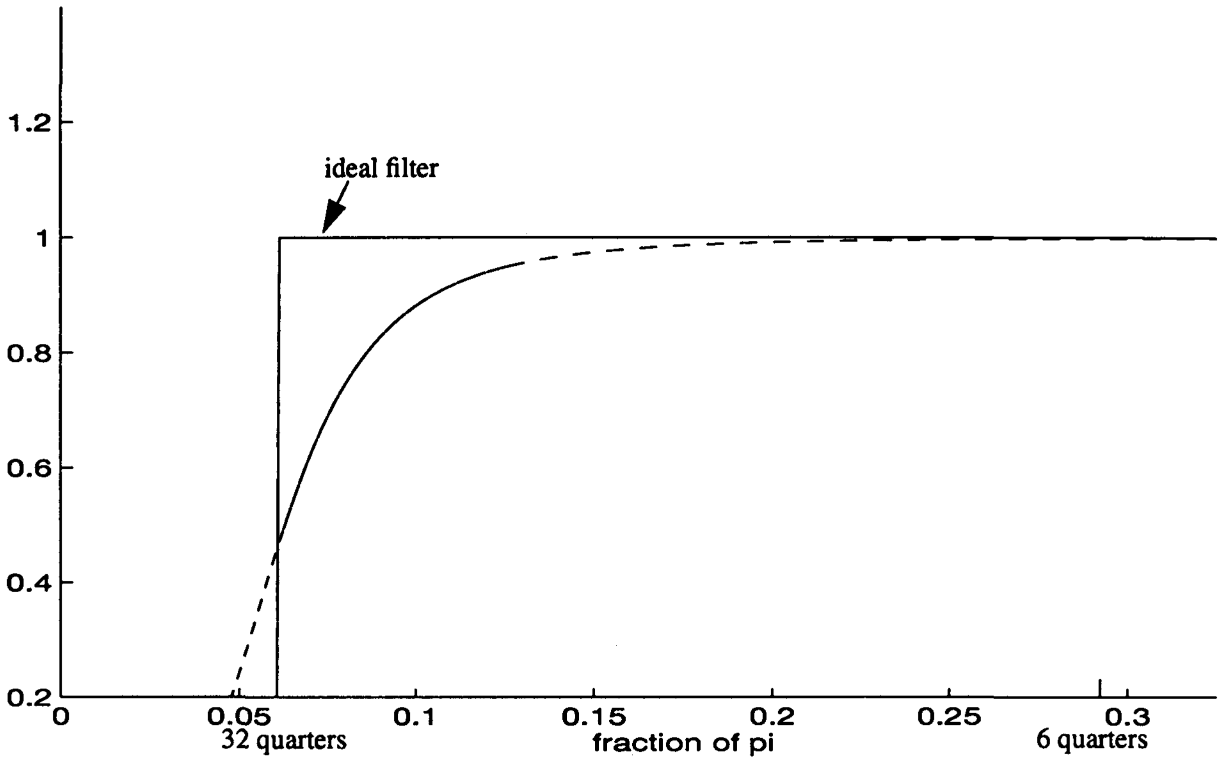
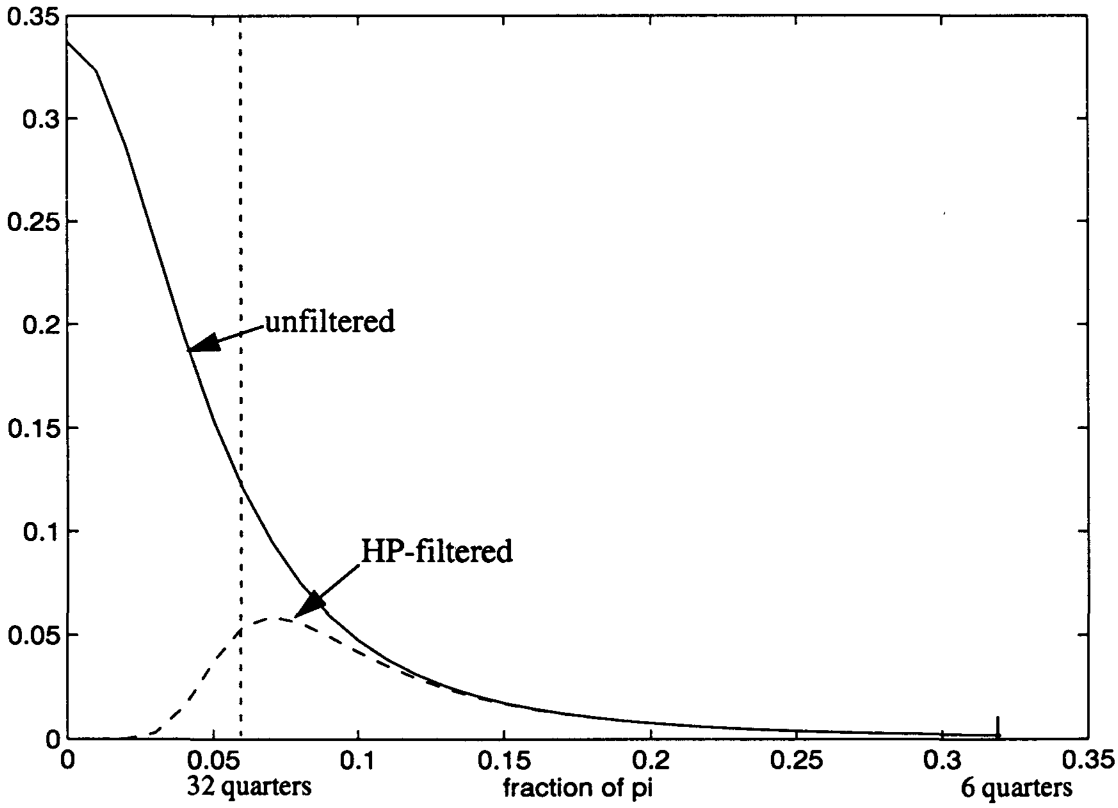


Figure 2

Series with the typical Granger shape - (AR(3) coefficients: 1.2, -0.11, -0.16)



One could associate the output gap with business cycle frequency plus higher frequency volatility in the data. Figure 1 would then suggest that the HP filter is an adequate measure of the output gap. One problem with this is that most macroeconomic time series are either integrated or highly persistent processes. In their study, Guay and St-Amant (1996) conduct a systematic investigation of the HP filter's ability to capture business-cycle frequencies; i.e. the area delimited by the spectrum of an original series at frequencies between 6 and 32 quarters. Their main finding is that, when the peak of a series is at zero frequency and the bulk of the variance is located in low frequencies, which is the shape described by Granger as typical for macroeconomic time series, the HP filter cannot capture business cycle frequencies adequately. This is illustrated by Figure 2, which shows the spectrum of an autoregressive process having its peak at zero frequency and that of the cyclical component resulting from the application of the HP filter.

In Figure 2, the spectrum of the cyclical component resulting from the application of the HP filter is very different from that of the original series. This comes as no surprise since the filter is designed to extract low frequencies from the data. However, we can see that business cycle frequencies are not left intact. In particular, the HP filter induces a peak inside business-cycle frequencies even though it is absent from the original series. Moreover, it fails to capture a significant fraction of the variance contained in business-cycle frequencies but captures some variance originating outside these frequencies. Guay and St-Amant (1996) show that this is typical of time series having the typical Granger shape; i.e. most macroeconomic series. Indeed, the unfiltered spectrum shown in Figure 2 is a parametric estimate of the spectrum of US real GDP.

The intuition behind this result is simple. Figure 1 shows that the gain of the HP filter at low business-cycle frequencies is smaller than that of the ideal filter. Indeed, the squared gain of the HP filter is 0.49 at frequencies corresponding to 32-quarter cycles and does not reach 0.95 before frequency $\pi/8$ (cycles of 16 quarters). Note also that the squared gain does not fall immediately to zero at lower frequencies. The problem is that a large fraction of the power of typical macroeconomic time series is concentrated in the band where the squared gain of the HP filter differs from that of an ideal filter. Also, the shape of the squared gain of the HP filter is such that when it is applied to typical macroeconomic time series a peak in the spectrum of the cyclical component is induced. In short, applying the HP filter to series dominated by low frequencies results in the extraction of a cyclical component that does not capture an important fraction of the variance contained in business-cycle frequencies of the original series captures an important part of the variance situated at lower frequencies than business-cycle frequencies but induces spurious dynamic properties.

An additional problem is that associating the output gap with the business-cycle frequencies in the data might not be a good idea in the first place. Note in particular that part of the variance associated with business-cycle frequencies could reflect the dynamics of shocks to potential output. As noted by King et al. (1991), "productivity shocks set off transitional dynamics, as capital is accumulated and the economy moves towards a new steady-state". To the extent that such dynamics reflect the evolution of potential output itself, one might prefer to use a different approach to identify potential output and the output gap. Section 3 of the paper provides a more detailed discussion of this point.

1.3 How well does the HP filter extract the cyclical component?

In the previous section, we have seen that the HP filter does not have spectral properties good enough to be able to isolate accurately the component of a series due to fluctuations at business cycle frequencies. As discussed by King and Rebelo (1993), another justification for the use of the HP filter is that in some cases it will be the optimal filter for identifying the cyclical component of a series. However, (King and Rebelo op. cit., and Harvey and Jaeger (1993)) these are cases when, in particular, the series is I(2), there are identical propagation mechanisms for innovations in the growth rate and in the cycle (or the transitory component is white noise), and the smoothing parameter λ is known. These conditions are rarely met in practice.

Of course, the fact that the HP filter is not an optimal filter does not necessarily mean that it will not be a good approximation of the optimal filter. We therefore turn to consider whether the HP filter can reliably isolate the cyclical component of a variety of time series.

It is often argued that macroeconomic time series are really comprised of a permanent component and a cyclical component. The permanent component could be driven by an I(1) technological process with drift, while monetary shocks, among others, could generate the cyclical component. In order to assess the HP filter's ability to extract such a cyclical component, consider the following DGP:

$$y_t = \mu_t + c_t, \text{ where} \tag{6}$$

$$\mu_t = \mu_{t-1} + \varepsilon_t \tag{7}$$

$$c_t = \phi_1 c_{t-1} + \phi_2 c_{t-2} + \eta_t \text{ and} \tag{8}$$

$$\varepsilon_t \sim NID\left(0, \sigma_\varepsilon^2\right), \quad \eta_t \sim NID\left(0, \sigma_\eta^2\right) \tag{9}$$

Equation (6) defines y_t as the sum of a permanent component, μ_t, which in this case corresponds to a random walk, and a cyclical component, c_t.[7] The dynamics of the cyclical component are specified as a second order autoregressive process so that the peak of the spectrum could be at zero frequency or at business-cycle frequencies. We assume that ε_t and η_t are uncorrelated.

Data are generated from equation (6) with ϕ_1 set at 1.2 and different values for ϕ_2 to control the location of the peak in the spectrum of the cyclical component. We also vary the standard-error ratio for the disturbances $\sigma_\varepsilon/\sigma_\eta$ to change the relative importance of each component. We follow the standard practice of giving the value 1,600 to λ, the HP filter smoothness parameter. We also follow Baxter and King's (1995) suggestion of dropping 12 observations at the beginning and at the end of the sample which should favour the filter considerably by abstracting partly from its end-of-sample problems (see Section 1.4). The resulting series contains 150 observations, a standard size for quarterly macroeconomic data. The number of replications is 500.

The performance of the HP filter is assessed by comparing the autocorrelation function of the cyclical component of the true process with that obtained from the filtered data. We also calculate the correlation between the true cyclical component and the filtered cyclical component and report their relative standard deviations $\left(\hat{\sigma}_c / \sigma_c\right)$. Table 1 presents the results and illustrates that the HP filter performs particularly poorly when there is an important permanent component. Indeed, in most cases, for high $\sigma_\varepsilon/\sigma_\eta$ ratios, the correlation between the true and the filtered components is not significantly different from zero. The estimated autocorrelation function is invariant to the change in the cyclical component in these cases (the values of the true autocorrelation functions are given in parentheses.) When the ratio $\sigma_\varepsilon/\sigma_\eta$ is equal to 0.5 or 1 and the peak of the cyclical component is located at zero frequency ($\phi_2 < -0.43$), the dynamic properties of the true and the filtered cyclical components are significantly different, as indicated by the estimated parameter values. In general, the HP filter adequately characterizes the series' dynamics when the peak of the spectrum is at business-cycle frequencies and the ratio $\sigma_\varepsilon/\sigma_\eta$ is small. However, even when the ratio of standard deviations is equal to 0.01 (i.e. the permanent component is almost absent), the filter performs poorly when the peak of the spectrum of the cyclical component is at zero frequency. Indeed, for $\phi_2 = -0.25$, the dynamic properties of the filtered component differ significantly from those of the true cyclical component, the correlation is only equal to 0.66, and the standard deviation of the filtered cyclical component is half that of the true cyclical component.

---

7 This is Watson's (1986) specification for real GDP in the United States.

Table 1
Simulation results for the HP filter

GDP			Estimated values				
			Autocorrelations				
$\sigma_\varepsilon/\sigma_\eta$	ϕ_1	ϕ_2	1	2	3	correlation	$(\hat{\sigma}_c / \sigma_c)$
10	0	0	0.71[0]	0.46[0]	0.26[0]	0.08	12.96
			(0.59, 0.80)	(0.30, 0.60)	(0.08, 0.43)	(-0.07,0.21)	(10.57,15.90)
10	1.2	-0.25	0.71[0.96]	0.47[0.90]	0.27[0.84]	0.08	4.19
			(0.61, 0.80)	(0.31, 0.61)	(0.08, 0.44)	(-0.11, 0.28)	(2.77, 6.01)
10	1.2	-0.40	0.71[0.86]	0.46[0.63]	0.26[0.41]	0.13	6.34
			(0.60, 0.80)	(0.30, 0.60)	(0.08, 0.44)	(-0.12, 0.36)	(4.82, 8.07)
10	1.2	-0.55	0.71[0.77]	0.46[0.38]	0.26[0.03]	0.14	6.93
			(0.60, 0.80)	(0.29, 0.60)	(0.06, 0.43)	(-0.08, 0.33)	(5.36, 8.70)
10	1.2	-0.75	0.71 [0.69]	0.46[0.27]	0.25[-0.19]	0.15	6.37
			(0.60, 0.78)	(0.30, 0.59)	(0.07, 0.41)	(-0.01, 0.31)	(4.79, 7.95)
5	0	0	0.69[0]	0.45[0]	0.26[0]	0.15	6.5
			(0.58, 0.78)	(0.30, 0.58)	(0.09, 0.41)	(0.02, 0.27)	(5.28, 7.85)
5	1.2	-0.25	0.71[0.96]	0.46[0.90]	0.26[0.84]	0.16	2.11
			(0.61, 0.80)	(0.32, 0.61)	(0.08, 0.43)	(-0.01, 0.36)	(1.43, 3.04)
5	1.2	-0.40	0.72[0.86]	0.46[0.63]	0.25[0.41]	0.23	3.26
			(0.61, 0.80)	(0.31, 0.60)	(0.08, 0.42)	(-0.01, 0.45)	(2.47, 4.15)
5	1.2	-0.55	0.71[0.77]	0.46[0.38]	0.24[0.03]	0.24	3.60
			(0.61, 0.80)	(0.30, 0.59)	(0.06, 0.41)	(0.01, 0.44)	(2.83, 4.52)
5	1.2	-0.75	0.70[0.69]	0.43[0.27]	0.20[-0.19]	0.29	3.3
			(0.61, 0.79)	(0.26, 0.57)	(0.00, 0.38)	(0.11, 0.44)	(2.53, 4.17)
1	0	0	0.43[0]	0.28[0]	0.20[0]	0.59	1.61
			(0.27, 0.57)	(0.11, 0.42)	(-0.02, 0.31)	(0.49, 0.70)	(1.41, 1.85)
1	1.2	-0.25	0.76[0.96]	0.51[0.90]	0.29[0.84]	0.51	0.66
			(0.67, 0.83)	(0.37, 0.62)	(0.11, 0.44)	(0.33, 0.68)	(0.44, 0.91)
1	1.2	-0.40	0.75[0.86]	0.44[0.63]	0.16[0.41]	0.71	1.02
			(0.67, 0.81)	(0.28, 0.55)	(-0.03,0.33)	(0.56, 0.82)	(0.83, 1.22)
1	1.2	-0.55	0.72[0.77]	0.34[0.38]	0.01[0.03]	0.76	1.15
			(0.66, 0.78)	(0.21, 0.47)	(-0.17,0.19)	(0.56, 0.82)	(0.83, 1.22)
1	1.2	-0.75	0.68[0.69]	0.15[0.27]	-0.27[-0.19]	0.83	1.16
			(0.63, 0.72)	(0.04, 0.27)	(-0.44,0.10)	(0.75, 0.89)	(1.04, 1.29)
0.5	0	0	0.16[0]	0.10[0]	0.04[0]	0.82	1.16
			(0.01, 0.32)	(-0.04, 0.24)	(-0.10, 0.18)	(0.75, 0.88)	(1.07, 1.27)
0.5	1.2	-0.25	0.79[0.96]	0.53[0.90]	0.30[0.84]	0.61	0.55
			(0.71, 0.85)	(0.38, 0.65)	(0.11, 0.46)	(0.41, 0.79)	(0.37, 0.76)
0.5	1.2	-0.40	0.77[0.86]	0.43[0.63]	0.13[0.41]	0.84	0.87
			(0.69, 0.81)	(0.29, 0.54)	(-0.05, 0.29)	(0.73, 0.92)	(0.74, 0.99)
0.5	1.2	-0.55	0.72[0.77]	0.28[0.38]	-0.10[0.03]	0.89	0.98
			(0.67, 0.78)	(0.17, 0.39)	(-0.25,0.06)	(0.83, 0.94)	(0.89, 1.07)
0.5	1.2	-0.75	0.67[0.69]	0.07[0.27]	-0.42[-0.19]	0.94	1.02
			(0.63, 0.71)	(-0.03, 0.18)	(-0.57, -0.27)	(0.90, 0.96)	(0.97, 1.08)
0.01	0	0	-0.08[0]	-0.06[0]	-0.06[0]	0.98	0.97
			(-0.21, 0.06)	(-0.21, 0.06)	(-0.19, 0.06)	(0.96, 0.99)	(0.94, 0.99)
0.01	1.2	-0.25	0.80[0.96]	0.54[0.90]	0.30[0.84]	0.66	0.51
			(0.72, 0.86)	(0.38, 0.67)	(0.11, 0.48)	(0.45, 0.83)	(0.34, 0.69)
0.01	1.2	-0.40	0.78[0.86]	0.43[0.63]	12[0.41]	0.90	0.81
			(0.72, 0.83)	(0.30, 0.55)	(-0.05, 0.28)	(0.82, 0.96)	(0.71, 0.90)
0.01	1.2	-0.55	0.73[0.77]	0.26[0.38]	-0.14[0.03]	0.96	0.92
			(0.67, 0.77)	(0.15, 0.37)	(-0.30, 0.01)	(0.91, 0.99)	(0.86, 0.96)
0.01	1.2	-0.75	0.67[0.69]	0.02[0.27]	-0.50[-0.19]	0.99	0.97
			(0.62, 0.71)	(-0.08, 0.13)	(-0.61, -0.35)	(0.97, 1.0)	(0.95, 0.99)

It is interesting to note that the HP filter does relatively well when the ratio $\sigma_\varepsilon/\sigma_\eta$ is equal to 1, 0.5, or 0.01 and the spectrum of the original series has a peak at zero frequency and at business-cycle frequencies (i.e. the latter frequencies contain a significant part of the variance of the series). Consequently, the conditions required to adequately identify the cyclical component with the HP filter can be expressed in the following way: the spectrum of the original series must have a peak located at business-cycle frequencies, which must account for an important part of the variance of the series. If the variance of the series is dominated by low frequencies, which is the case for most macroeconomic series in levels, including real output, the HP filter does a poor job of extracting an output gap associated with the cyclical component of real output.

1.4 The HP filter at the end of samples

In examining the performance of the HP filter in the last two sections, we have looked at how well it isolates particular business cycle frequencies or the cyclical component of the series. Both cases implicitly looked at the performance of the HP filter over the available sample of data as a whole. However, it is useful to remember that the focus for policy advice is on estimating the *current* output gap. This is a more difficult task since future information will presumably be useful in determining whether recent changes in output are persistent or transitory. We should therefore consider how the conclusions from the two previous sections might be altered by this added complication.[8]

To understand how the HP filter behaves at the end of sample, recall that the optimization problem it solves trades off the size of deviations from trend with the smoothness of that trend. In the face of a transitory shock, the filter is therefore reluctant to change the trend very much since this implies raising the trend before the shock and lowering it afterwards. However, the latter penalty is absent, implying that the optimal trend will be more responsive to transitory shocks than in mid-sample.

We can show this difference in several ways. Figure 3 shows the HP filter trend expressed as a moving average of the unfiltered data. The weights in this moving average change as we move from the mid-sample towards the end of sample. The former gives us a smooth 2-sided average in which no observation receives more than 6% of the weight. However, the latter gives a 1-sided average where the last observation alone accounts for 20% of the weight. Not surprisingly, this makes the HP trend more variable at the end of sample. Figure 4 shows that the deviations from the HP trend result in different frequency responses. In particular, the 1-sided, or end-of-sample, filtered deviations from trend capture less of the variation at business cycle frequencies (indicated by the dotted vertical lines).[9]

Figure 5 and Figure 6 show how the deviations from the HP trend differ depending on whether we are at the end-of-sample or mid-sample. The solid line in Figure 5 shows the usual deviation from the HP trend for Canadian GDP. The dashed line then shows the estimate we get from the same if we only use data available up to that point in time (i.e. the corresponding end-of-sample estimate). Although the two series tend to move together, there are some important differences in size and timing. Comparing Figure 5 with Figure 6, we see that while deviations from trend are usually less than 3% of GDP, the difference between its mid-sample and end-of-sample estimates is often as large as 2% of GDP. If we accept that the difference between these two measures is just one

---

8 This problem has been mentioned in other studies as well. Much of the analysis we present can also be found in Butler (1996).

9 The squared-gains of the two HP trends also look quite different. At the frequency corresponding to cycles of 32 quarters, the end-of-sample filter has a squared-gain of about 1 while the mid-sample filter has a squared-gain of about 0.1. This is precisely the frequency at which an optimal filter would have a gain of zero.

Figure 3

MA representation of the HP filter as a function of sample position (128 observations, λ = 1,600)

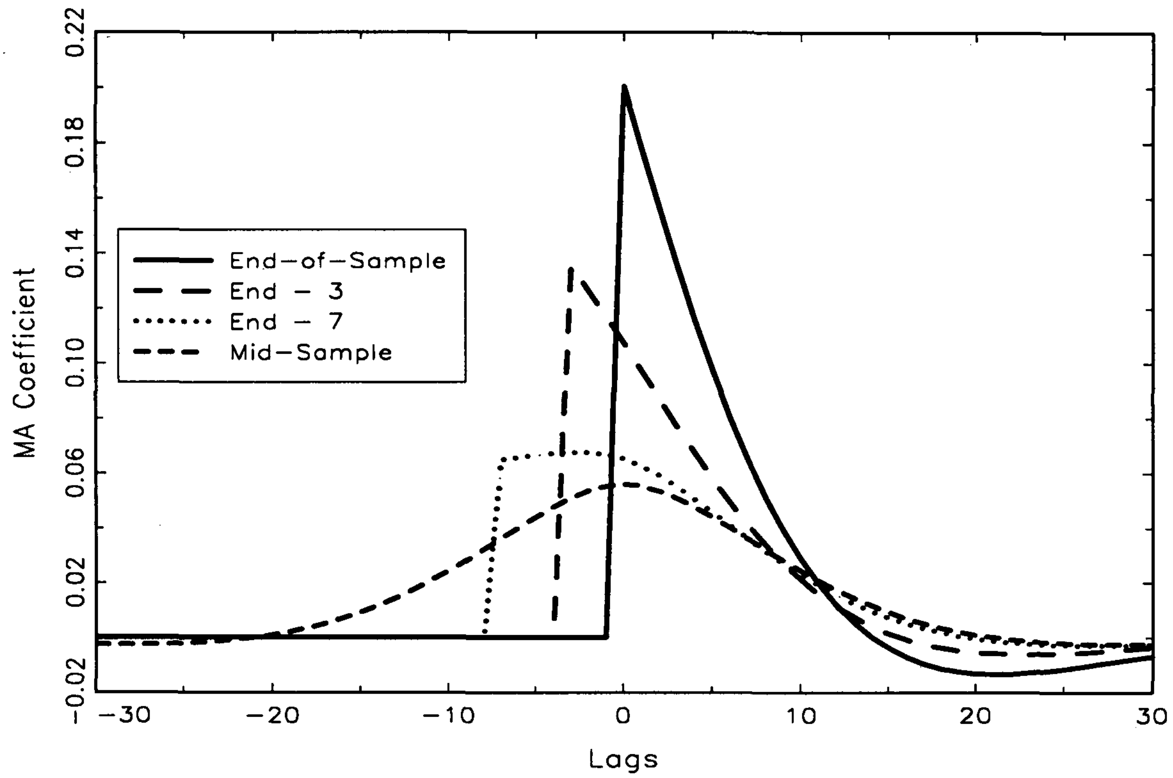


Figure 4

Squared gain of the HP filter (128 observations, λ = 1,600)

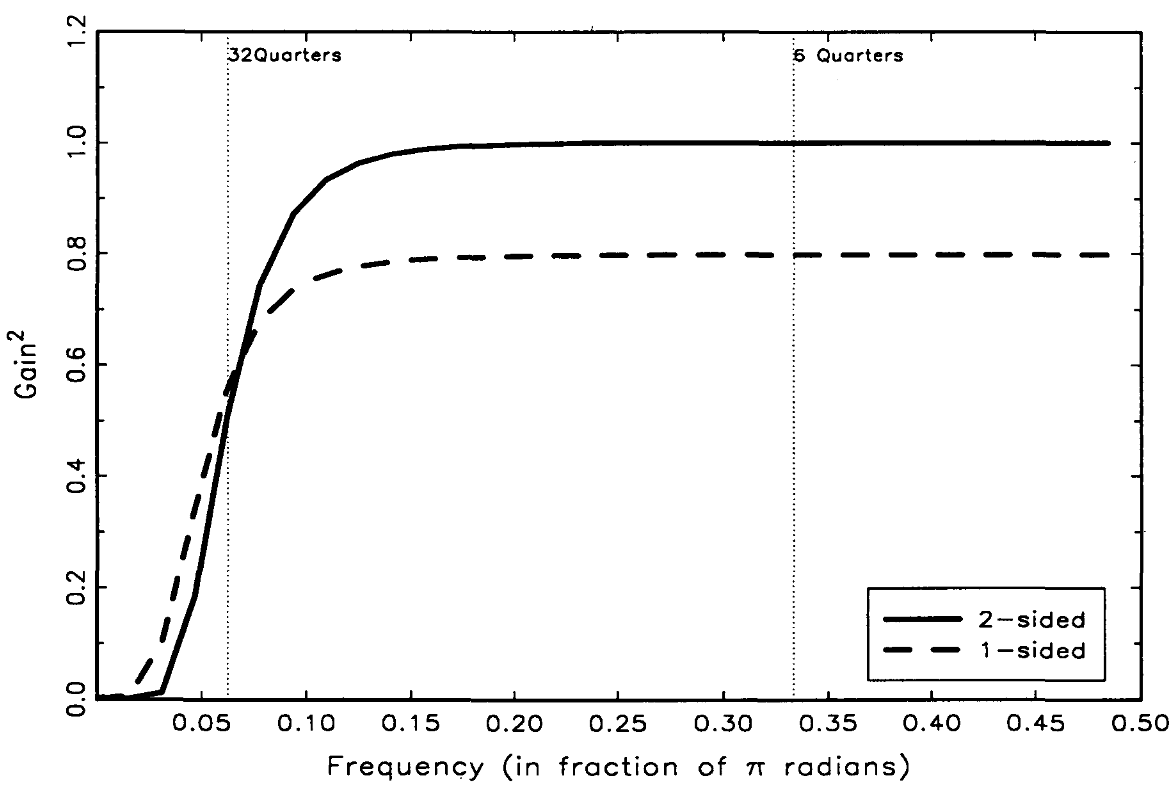


Figure 5

HP detrended real GDP (Canada, λ = 1,600)

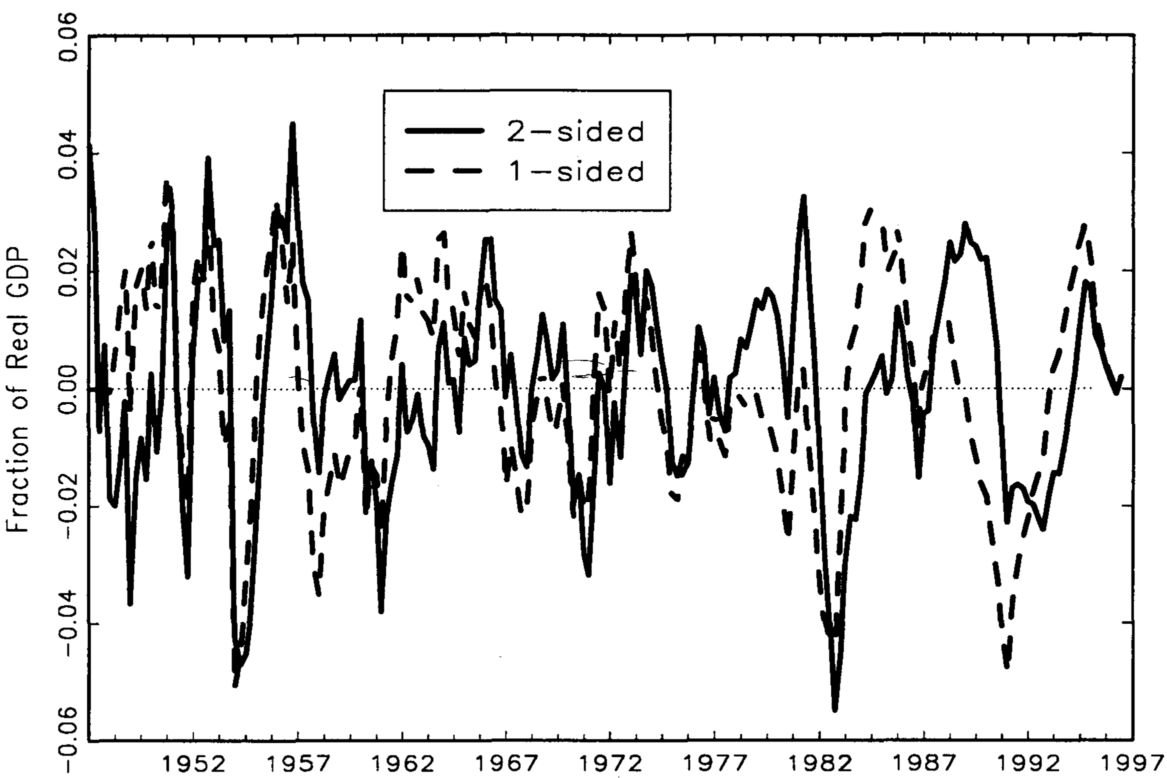


Figure 6

HP detrended real GDP mid-sample – end-of-sample (Canada, λ = 1,600)

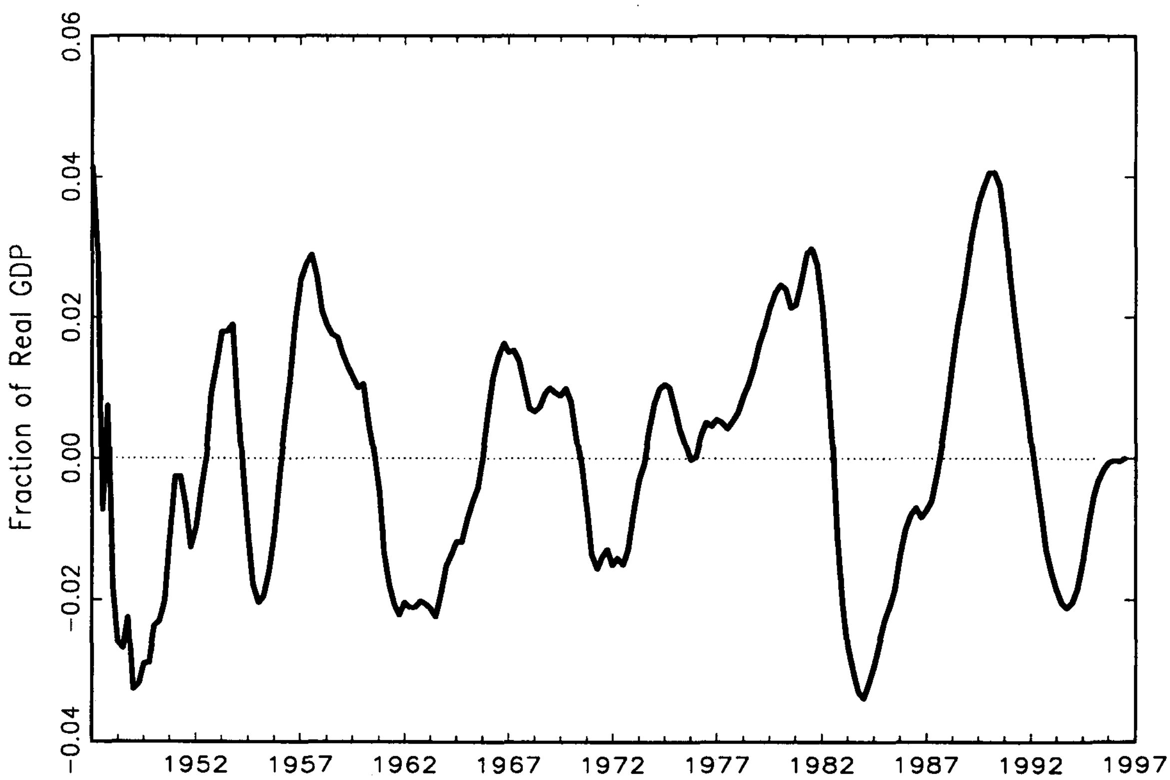
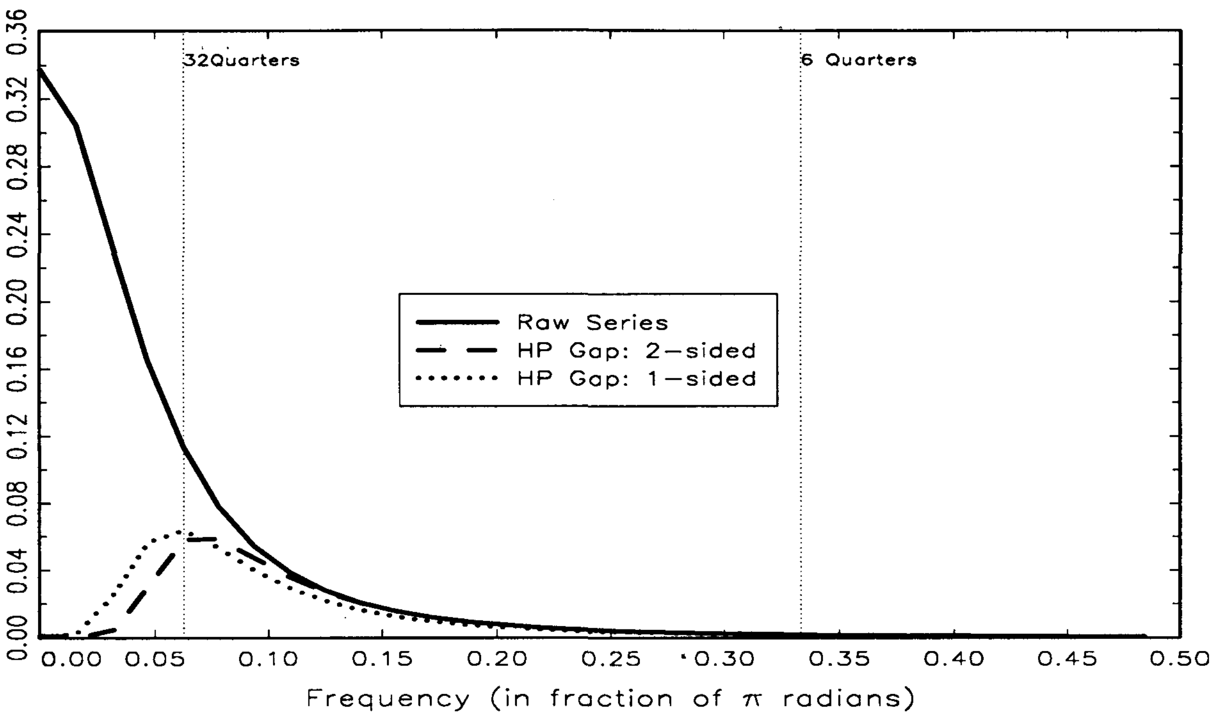


Figure 7

Spectrum of series with typical Granger shape (128 observations, λ = 1,600)



component of the measurement error of end-of-sample estimates, then the measurement errors of the latter must be roughly as large as the estimates themselves.[10] Hence, end-of-sample estimates cannot be very reliable estimates of deviations from trend.

Figure 7 applies the HP filter to the "typical Granger Shape" series we considered previously. At the end-of-sample, even less of the variance of the deviations from the HP trend is due to variations at business cycle frequencies and more is due to "leakage" from lower frequencies. This suggests that the results we obtained in Section 1.2 probably overstate the reliability of the HP filter for identifying an output gap associated with business cycle frequencies. This is consistent with the results of Laxton and Tetlow (1992) and Butler (1996), who note that related filters also seem to perform worse at the end of samples. We turn to these related filters in Section 2.

1.5 Limits to 1-sided filtering

Part of the end-of-sample problem discussed in Section 1.4 reflects the fact that the HP filter behaves differently at the end-of-sample and at mid-sample, as shown in Figure 3. This suggests that other univariate filters might be able to measure output gaps more reliably. In this section, we consider one intrinsic limit to the ability of univariate filters to measure the current output gap, and show how this in turn will relate to beliefs about the economic relationships between actual and potential output. We show that models in which potential output is exogenous with respect to actual

[10] We reach the same conclusion if we look at the range of the series, or at their standard deviations. The range (maximum - minimum) of the 1-sided estimate is 8.7% of GDP while the range of the difference between the 1 and 2-sided estimates is 7.5%; the comparable standard errors are 1.8% and 1.8%. These comparisons are only approximate; small sample problems in the 1-sided estimate at the beginning of the sample may make their difference appear excessively volatile, while constraining the two estimates to be identical at the end of the sample will tend to understate the volatility of their difference.

output and the output gap imply that univariate filters will never be able to give much information about contemporaneous output gaps.[11]

Suppose that potential output can be expressed as a linear filter of actual output, so that:

$$q_t = A(L).\, y_t + \varepsilon_t \tag{10}$$

where q_t is (the log of) potential output, y_t is (the log of) actual output, ε_t is an innovations process that is uncorrelated with y_t at all leads and lags, and $A(L)$ is a two-sided polynomial in the lag operator (i.e. it takes a weighted sum of leads, lags and contemporaneous values of y_t). A sufficient but not necessary condition for such a representation to exist is that output y_t has a unit root and that the output gap $q_t - y_t$ is stationary.

We typically think of y_t as being non-stationary in mean, since it tends to drift upwards over time. To ensure that q_t and y_t move together in the long run (so that the gap is stationary), we will further assume that:

$$A(1) = 1 \tag{11}$$

which simply means that the weights in $A(L)$ must sum to one. This in turn implies that:

$$A(L) - 1 = (1 - L).\, \tilde{A}(L), \text{ and therefore that:} \tag{12}$$

$$q_t - y_t = \tilde{A}(L).\, \Delta y_t + \varepsilon_t \tag{13}$$

Therefore, we should be able to express the output gap as the weighted sum of past, present and future output growth. The difference between equation (13) and the HP filter representation of the output gap is that the HP filter implies a particular set of restrictions on $\tilde{A}(L)$ that vary with the position in the sample. Let:

$$E\left(q_t - y_t \middle| H_y\right) = \tilde{A}(L).\, \Delta y_t \text{ and} \tag{14}$$

$$V\left((q_t - y_t) - E\left(q_t - y_t \middle| H_y\right)\right) = \sigma^2 \tag{15}$$

where H_y is the set of all past, present and future values of y_t.

So far, we have assumed that $\tilde{A}(L)$ is two-sided, whereas its use for policy purposes requires that it be one-sided. To understand how such a restriction on $\tilde{A}(L)$ will affect the accuracy of our estimate, note that the law of iterated expectations and equation (14) imply:

$$E\left(q_t - y_t \middle| H_y^-\right) = E\left(E\left(q_t - y_t \middle| H_y\right)\middle| H_y^-\right) = E\left(\tilde{A}(L).\, \Delta y_t \middle| H_y^-\right) \tag{16}$$

where H_y^- is the set of all past values of y_t. If we define:

$$\tilde{A}(L) = \tilde{A}^-(L) + \tilde{A}^+(L) \tag{17}$$

where $\tilde{A}^-(L)$ has only positive powers of L and $\tilde{A}^+(L)$ only non-positive powers, then equation (16) implies:

$$E\left(q_t - y_t \middle| H_y^-\right) = E\left(\tilde{A}^-(L).\, \Delta y_t \middle| H_y^-\right) + E\left(\tilde{A}^+(L).\, \Delta y_t \middle| H_y^-\right) \tag{18}$$

$$= \tilde{A}^-(L).\, \Delta y_t + \sum_{j=1}^{m} a_j^+ .\, E\left(\Delta y_{t+j} \middle| H_y^-\right) = \bar{A}(L).\, \Delta y_t$$

[11] This section draws heavily on van Norden (1995).

where a_j^+ is simply the coefficient on L^{-j} in $\tilde{A}^+(L)$. Similarly, we can show that

$$V\left(q_t - y_t \middle| H_y^-\right) = V\left(q_t - y_t \middle| H_y\right) + V\left(\tilde{A}(L).\Delta y_t \middle| H_y^-\right) = \sigma^2 + V\left(\tilde{A}^+(L).\Delta y_t \middle| H_y^-\right) \qquad (19)$$

where $V(X|\Omega)$ is the variance of the error in forecasting X given the information set Ω.

Equations (18) and (19) have an intuitive interpretation. The extent to which the 1-sided filter $\overline{A}(L)$ is less informative than $\tilde{A}(L)$ will depend on the weight which $\tilde{A}(L)$ puts on current and future values of Δy_t and the extent to which those future values can be predicted from current and past values. The former will in turn depend on the Granger causal relationship between $q_t - y_t$ and Δy_t, while the latter will depend on the degree to which output growth is serially correlated.

For most industrialized nations, 12 lags of quarterly output growth predicts only 20-40% of the variance of current output growth and much of this explanatory power seems to come from the first few lags. This suggest that since predictability can be low, the extent of Granger-causality will play an important role in determining how accurately the one-sided univariate filter can estimate the current output gap. For simplicity, we will discuss the role Granger-causality plays under the assumption that the past history of output growth is of no use in predicting present and future output growth.

From equation (18) we can see that $\overline{A}(L)$ will tell us as much about the output gap as $\tilde{A}(L)$ when $\tilde{A}^+(L) = 0$, which in turn implies that $q_t - y_t$ does not Granger-cause Δy_t. Van Norden (1995) shows that the latter condition in turn implies that q_t does not Granger-cause y_t. If y_t and q_t are cointegrated, this would imply that there is unidirectional causality from y_t to q_t. In other words, exogenous shocks to potential output would have no subsequent effect on actual output, but persistent shocks to actual output would eventually be followed by a similar change in potential. Such behaviour could describe a particularly severe form of hysteresis; one where output has no tendency to return to potential and, instead. potential output is driven in the long run only by previous variations in actual output. In this kind of world, univariate filters can hope to be as effective in estimating the current output gap as they are in estimating past output gaps.

The latter will not be the case when $q_t - y_t$ Granger-causes Δy_t. Therefore, so long as output growth appears to respond to some degree to past changes in the output gap, then univariate methods will estimate the current output gap less accurately than past output gaps. The intuition behind this result should be clear. Since future output growth will reflect the influence of the current output gap, one can gain information about the current gap by observing future growth.

Univariate methods will be of no use in estimating the current output gap when Δy_t not Granger cause $q_t - y_t$.[12] In other words, if faster, or slower, than normal output growth tends to have no subsequent effect on the size of the output gap, then time series estimates of the current gap will be as uninformative as possible. The intuition is similar to that above. Past output growth is the only information about the gap that we have; if it tells us nothing about the current gap, then our estimates will be unilluminating.

It is more difficult to say what kind of economic model will generate this kind of result since Granger-causality from Δy_t to $q_t - y_t$ does not directly correspond to any statement about

[12] Again, this conclusion assumes that past output growth is of no use in predicting future variations in output growth. As was mentioned earlier, the data show that there is some serial correlation in output growth, so time series methods would still have some explanatory power even in this case.

Granger-causality between y_t and q_t.[13] However, it is possible to give examples in which this result would hold. One simple case would be where:

$$\Delta y_t = \alpha(y_{t-1} - q_{t-1}) + u_t \quad \text{and} \quad q_t = q_{t-1} + v_t \tag{20}$$

Potential output follows a random walk that is independent of the behaviour of output. Actual output in turn is generated by a simple error-correction model, which ensures that actual and potential output move together in the long run. Such a model precisely satisfies the condition for no Granger causality from Δy_t to $q_t - y_t$.

Clearly, there is a range of models in which univariate time-series methods will be of little use at the end of sample. Furthermore, it is the short-run dynamics of potential and actual output which are critical in determining whether models belong to this class. This is not an empirically testable question since we cannot directly observe potential. However, we can try to ensure that our views on the determination of potential are consistent with the methods we use to measure it.

2. Extensions of the HP filter

The Bank of Canada has used various extensions of the HP filter to obtain measures of the output gap and help guide policy. These "hybrid" methods were developed in the 1990s to try to balance strengths and weakness of "structural" and "astructural" approaches to measuring the output gap for policy makers. The key papers explaining the justification and implementation of this approach are Laxton and Tetlow (1992) and Butler (1996). Work in a similar spirit has been pursued both at some of the Federal Reserve Banks (such as Kuttner (1994)) and at the OECD (see Giorno et al. (1995)).

To understand the contribution of these methods, one needs to appreciate the problems that these authors were trying to avoid. Laxton and Tetlow argue that there is insufficient knowledge about the true structural determinants of the supply side of the economy to make the purely structural approach practicable. At the same time, for policy purposes we need to distinguish between movements in output caused by supply shocks and those caused by demand shocks, whereas most astructural (time-series) models attempt to distinguish permanent and transitory components of output. As an alternative, they suggest a way of combining the two approaches which we refer to as the multivariate HP filter.[14]

As we explain in Section 2.1, this methodology consists of adding the residuals of a structural economic relationship to the minimization problem that the HP filter is seeking to solve. Section 2.2 discusses the production function variant of this methodology. In Section 2.3, we examine additional modifications introduced to the filter to improve its performance at the end of the sample. Section 2.4 looks at these approaches from a different perspective and relates them to both the methods of Section 1 and other methods that use additional structural relationships.

---

13 See van Norden (1995).

14 Laxton and Tetlow refer to their specific filter as "The Multivariate Filter (MVF)" and Butler refers to his as "The Extended Multivariate Filter (EMVF)". In this paper, we broadly refer to all multivariate extensions of the univariate Hodrick-Prescott filter as Multivariate HP Filters (MHPF), which include the MVF and EMVF as special cases. The method currently used to estimate Canadian potential output for the Bank's staff projection will also be referred to as the EMVF. The latter differs somewhat from the implementation described in Butler (1996), but is conceptually the same.

2.1 A multivariate HP filter

As noted previously, the original HP filter chooses the trend as the solution to:

$$\left\{y_t^g\right\}_{t=0}^{T+1} = \arg\min \sum_{t=1}^{T}\left(y_t - y_t^g\right)^2 + \lambda\left(\Delta^2 y_{t+1}^g\right)^2 \tag{21}$$

where $\Delta^2 y_{t+1}^g = \Delta.\Delta.y_{t+1}$ and $\Delta y_t = y_t - y_{t-1}$. The HP filter adds a term,

$$\left\{y_t^g\right\}_{t=0}^{T+1} = \arg\min \sum_{t=1}^{T}\left(y_t - y_t^g\right)^2 + \lambda_g\left(\Delta^2 y_{t+1}^g\right)^2 + \lambda_\varepsilon \varepsilon_t^2 \tag{22}$$

where $\varepsilon_t = z_t - f(y_t^g, x_t)$. z_t is some other economic variable of interest, and $f(.)$ models z_t as a function of both some explanatory variables x_t and the unobserved trend y_t^g. The new term in ε_t^2 has the effect of choosing the trend to simultaneously minimize deviations of output from trend, minimize changes in the trend's growth rate, *and* maximize the ability of the trend to fit some structural economic relationship $f(.)$. λ_g and λ_ε reflect the relative weights of these different objectives.

The key to implementing the multivariate HP filter for the purpose of estimating potential output (or an output gap) is to specify $(z_t, f(y_t^g, x_t))$ in such a way as to capture some structural relationship that depends on either potential output or the output gap. For example, one could specify a Phillips curve equation that relates observed inflation to a measure of inflation expectations, the output gap, and perhaps additional explanatory variables (such as oil prices). ε_t would then be the residual from this Phillips curve equation and trend output would be chosen in part to improve the explanatory power of the output gap for inflation. Alternatively, one could use an Okun's Law relationship to link the rate of unemployment to the output gap and various structural variables determining the natural rate of unemployment. The trend of output would then be influenced by the evolution of the unemployment rate and its structural determinants.

Of course, there is no reason for restricting ourselves to a single structural relationship. Equation (22) can be generalized to include an arbitrary number n of structural relationships with a common trend y_t^g, giving

$$\left\{y_t^g\right\}_{t=0}^{T+1} = \arg\min \sum_{t=1}^{T}\left(y_t - y_t^g\right)^2 + \lambda_g\left(\Delta^2 y_{t+1}^g\right)^2 + \left(\sum_{i=1}^{n} \lambda_{\varepsilon_i} \varepsilon_{it}^2\right) \tag{23}$$

The original Laxton and Tetlow (1992) paper used information from both a Phillips curve and an Okun's Law relationship and Butler (1996) also uses multiple structural relationships simultaneously.

The usefulness of the multivariate HP filter depends on several factors. Obviously, the extent to which it improves upon the original HP filter will depend on the reliability and information content of the structural relationship(s) with which it is combined. These potentially offer a way of mitigating the problems of HP filters noted in Section 1. However, given the importance of obtaining good end-of-sample estimates of output gaps, we require structural relationships that can give good contemporaneous information.[15]

For the particular DGP they examine, Laxton and Tetlow find that the degree to which their filter does better than the univariate HP filter at estimating the output gap increases with the relative importance of demand shocks to supply shocks. While the improvement produced by the

---

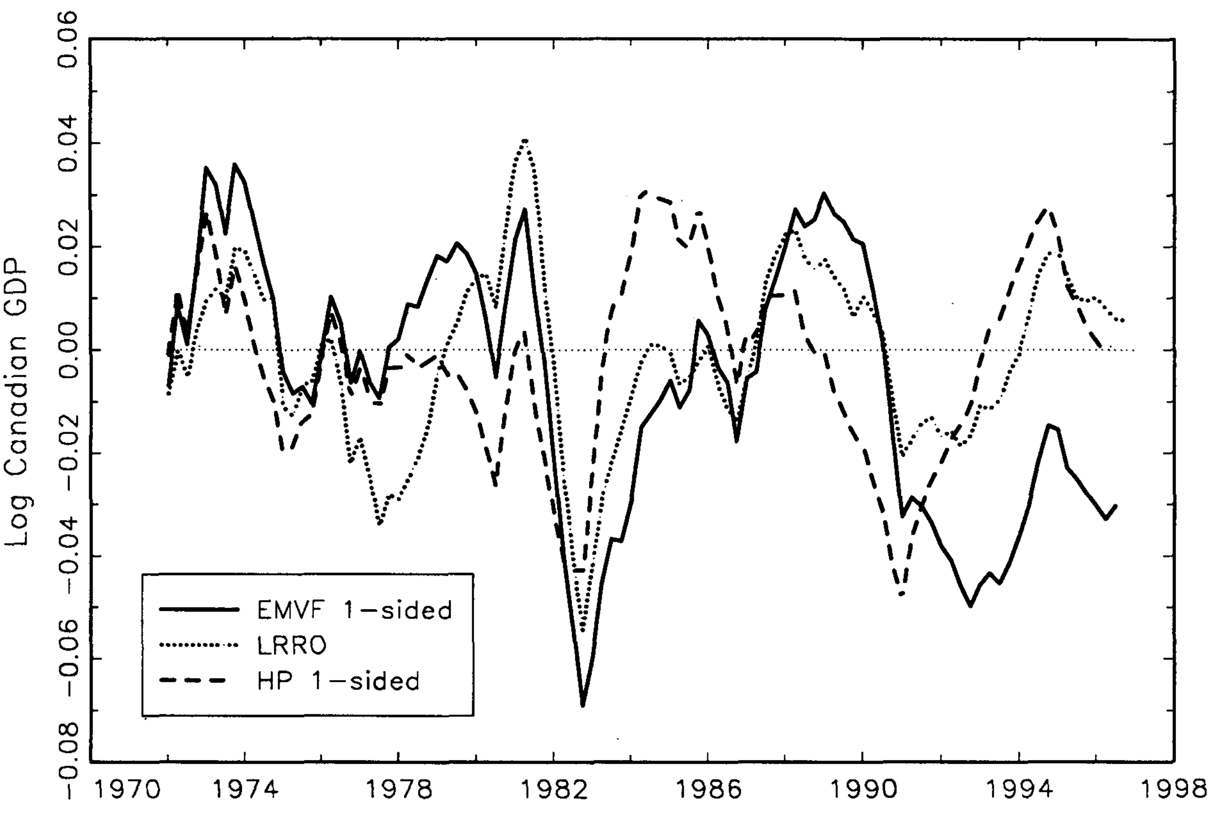
[15] In that respect, there may be limitations to the information we can expect to gain from Phillips curve relationships if we believe that inflation responds to output gaps with a lag.

MHPF can be large, they find that there is still substantial uncertainty in their point estimates of the output gap and that this uncertainty is larger at the end of sample. In their base case, they find that the 95% confidence interval for the output gap at the end of sample is about 4%, which implies that policymakers would rarely observe statistically significant output gaps.

Another factor of key importance to the success of the multivariate HP filter is calibration. Instead of having a single λ parameter with a consensus value of 1,600, we now have vectors of parameters $\bar{\lambda}_{g,\varepsilon}$ without a clear guide as to their appropriate values. In addition, we also need to estimate the form of the structural relationships involving potential output or the output gap. If we attempt to do this before calculating $\{y_t^g\}$, then we will be estimating a structural relationship which may be inconsistent with the values of $\{y_t^g\}$ the MHPF produces. Furthermore, theory will often not be a sufficient guide to allow us to tightly calibrate such a relationship. The approach used by Laxton and Tetlow (1992) and Butler (1996) is to experiment with alternative weightings to see which produce reasonable results and how sensitive the outcomes are to these choices.

Figure 8

Comparison of 3 different measures of the Canadian output gap



An alternative explored by Harvey and Jaeger (1993) and Côté and Hostland (1994) is to estimate the structural relationship simultaneously with $\{y_t^g\}$ and $\{\lambda_g, \bar{\lambda}_\varepsilon\}$via maximum likelihood methods.[16] Côté and Hostland found that the results can be sensitive to the specification of the

[16] Butler (1996) mentions that a direct maximum likelihood estimation was attempted, but that this did not produce reasonable results for the λ's.

structural relationships,[17] that the usefulness of the structural information vanishes when one considers only end-of-sample performance,[18] that the structural parameters cannot be estimated with much accuracy, and that maximization of the likelihood function was problematic.

To give some idea of how such filters perform in practice, Figure 8 compares three different estimates of the output (GDP) gap. The first is that produced by the Butler (1996) filter (labelled EMVF).[19] The others are those produced by a 1-sided HP (1600) filter and by the LRRO filter (which we introduce below, in Section 3). We can see from that figure that all three methods produce gaps of roughly the same amplitude, and that there is a tendency for the three series to rise and fall at similar times. While all three series show negative output gaps (i.e. excess supply) in the early 1990s, the LRRO and the HP show the economy returning to potential after a few years, while the EMVF shows large output gaps remaining through the end of the sample (1996Q3.) However, as we see in the next section, this last difference is more a reflection of the differences in the structural information used.

2.2 The production function approach

Another important feature of the EMVF filter is that rather than directly filtering output, output is decomposed into a number of components which are then individually filtered. This allows for a more direct link to sources of structural information as well as for an easier interpretation of the source of changes in the gap or potential.

The decomposition is based on an aggregate Cobb-Douglas constant-returns-to-scale production function:

$$Y = QN^{\alpha}K^{1-\alpha} \qquad (24)$$

where Q is total factor productivity, N is labour, K the capital stock, and α is the labour-output elasticity (as well as labour's share of income). With some algebra, we can show that:

$$\mu \equiv \partial Y / \partial N = \alpha Y / N \Rightarrow y = n + \mu - \alpha \qquad (25)$$

where lower-cases letter are the logs of the upper-case counterparts. This means that to estimating the trend in output, we instead estimate the trends in employment, the marginal product of labour and the labour-output elasticity and then sum them. One nice feature of the decomposition in equation (25) rather than equation (24) is that it allows us to avoid the problem of trying to estimate reliably the capital stock. We can then use the further decomposition that the log of total employment n is given by:

$$n = Pop + p + \log(1 - u) \qquad (26)$$

where Pop is the log of the working-age population, p is the log of the participation rate and u is the rate of unemployment.

---

[17] They find that specifying the dynamic relationship in levels or first differences has a large effect on the estimated values of λ_{ε}.

[18] Côté and Hostland approximate the behaviour of the 1-sided filter by using only the lags from the mid-sample representation of the HP filter. They also obtain much more useful results when they apply the 2-sided filter at the end-of-sample using forecast values for the required leads in the filter.

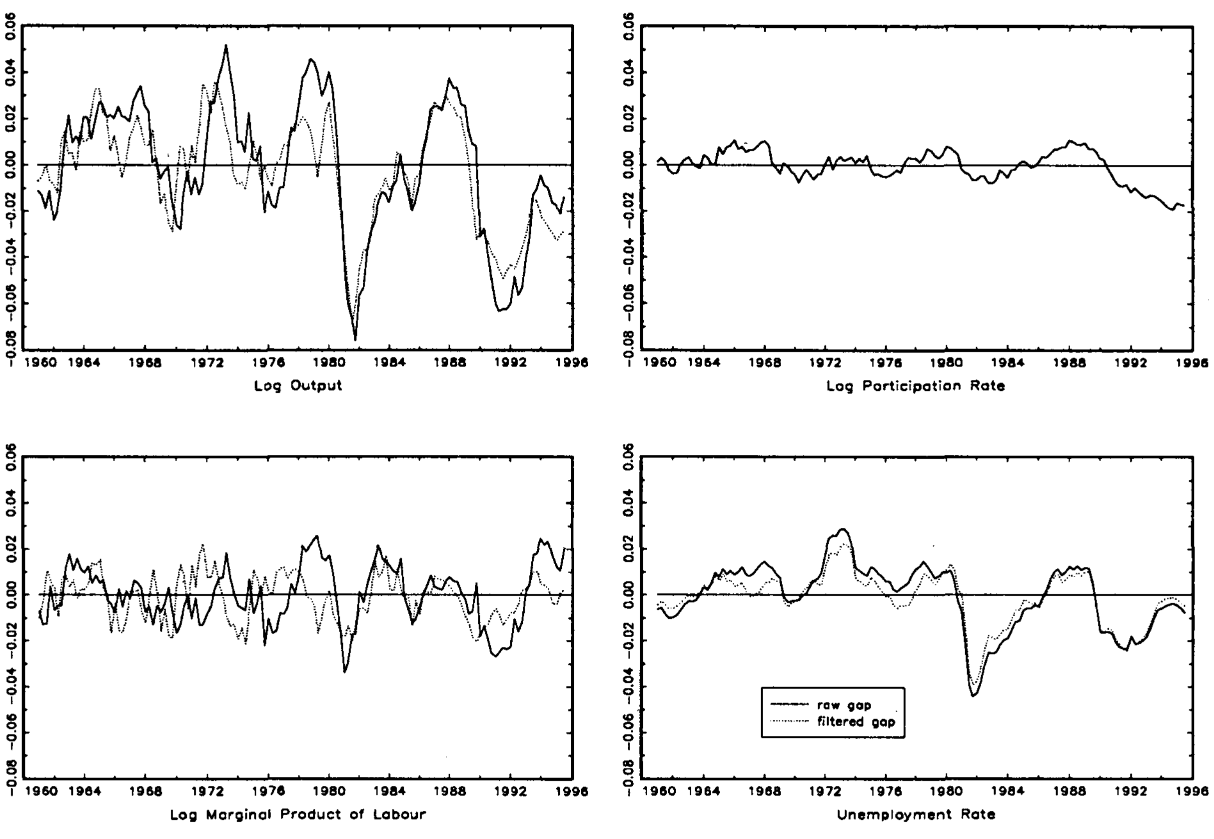
[19] The details of the filter used for the Bank's staff projection of the Canadian economy evolve over time in response to ongoing research and may therefore, as noted in footnote 14, differ slightly from the exposition in Butler (1996). The EMVF gaps shown in the figure reflect the specification used in the staff's December 1996 Economic Projection.

Within this framework, the level of potential output is defined to be the level of output consistent with existing population, trend rates of unemployment and participation and trend levels of the marginal product of labour and the output share of labour. In practice, the trend levels of the participation rate and the output share of labour are determined by a combination of judgement, demographic factors and univariate HP smoothing. Separate MHPF systems are then used to identify trend levels of the marginal product of labour and unemployment. The trend rate for unemployment is estimated with the help of a long-run structural equation due to Côté and Hostland (1996) as well as a price-unemployment Phillips Curve described in Laxton, Rose and Tetlow (1993). Perhaps more novel is the use of a long-run relationship between the marginal product of labour and producer wages as well as an Okun's Law relationship to identify the trend income share of labour.

Analysis of the performance of the EMVF in Butler (1996) shows that this method has its own strengths and weaknesses. On the one hand, he notes that the rolling and full-sample estimates of the NAIRU and the equilibrium marginal product of labour are quite similar, and that the labour market gaps are highly correlated with inflation. On the other hand, he also notes that there is significant correlation in the errors across structural equations, suggesting that there may be further efficiency gains to be had.

Figure 9

Decomposition of the EMVF output gap



To understand the source of the persistent output gap the EMVF produces in the 1990s, we can decompose the output gap into its three components, as shown by the dotted lines in Figure 9. This shows that the filter gives an unemployment rate gap and a marginal product of labour gap that are roughly zero at the start of 1996. Instead, the aggregate output gap largely reflects a deviation of the participation rate from its structural trend level. (Again, note that the participation rate gap is not filtered before it is added to the other components.) However, it would be wrong to attribute the aggregate gap entirely to structural information, as we can see by comparing the filtered output gap (top-right corner, dotted line) with its unfiltered, or "raw", counterpart (same graph, solid line). This shows that the effects of filtering over the most recent period has tended to increase the estimated size of the output gap by 1-2% of GDP. Since the filtered unemployment rate gap is very close to its unfiltered counterpart, this implies that most of the difference between the filtered and unfiltered output gap is due to the effects of the filter on the marginal product of labour gap.

2.3 The end-of-sample problem

Having seen that both the filter and the structural information play important roles in the EMVF's estimate of recent output gaps, nothing discussed to this point address the end-of-sample problems of HP filters noted in Section 1.4. However, the EMVF contains two novel features intended to modify its end-of-sample behaviour.

First, the EMVF contains an additional growth-rate restriction. If we temporarily ignore the structural information for expositional simplicity, the modified filter solves the problem:

$$\left\{y_t^g\right\}_{t=0}^{T+1} = \arg\min\left(\left(\sum_{t=1}^{T}\left(y_t - y_t^g\right)^2 + \lambda_g\left(\Delta^2 y_{t+1}^g\right)^2\right) + \sum_{t=T-j}^{T}\lambda_{ss}\left(\Delta\tau_t - \mu_{ss}\right)\right) \tag{27}$$

where μ_{ss} is a constant equal to the steady-state growth rate of potential output and λ_{ss} is the weight put on the growth rate restriction. The key feature is that λ_{ss} only penalizes deviations from the steady-state growth rate in the last j periods of the sample, effectively "stiffening" the filter. This restriction assumes that the growth rate of potential reverts toward a constant, whereas the theoretical justification for HP filters as optimal filters (noted in Section 1) assumes that this growth rate contains a stochastic trend and therefore will not show any such reversion. Whether such a restriction leads to more accurate estimates of the output gap depends on the accuracy with which one can determine the appropriate value of μ_{ss}.

The second novel feature of the EMVF's treatment of end-of-sample problems is the introduction of a recursive updating restriction. This simply adds an additional term to equation (27), giving:

$$\left\{y_t^g\right\}_{t=0}^{T+1} = \arg\min\left(\left(\sum_{t=1}^{T}\left(y_t - y_t^g\right)^2 + \lambda_g\left(\Delta^2 y_{t+1}^g\right)^2\right) + \lambda_{pr}\left(y_t^g - {}_{pr}y_t^g\right)^2 + \sum_{t=T-j}^{T}\lambda_{ss}\left(\Delta\tau_t - \mu_{ss}\right)\right) \tag{28}$$

where ${}_{pr}y_t^g$ is the tth element of $\{y_t^g\}_{t=0}^{T}$. This means that filter is restricted to choose $\{y_t^g\}$ while minimizing the degree to which new observations modify estimates of y^g based on shorter spans of observations. Perhaps not surprisingly, Butler (1996) shows that this gives a 1-sided estimate of the output gap that behaves more like the subsequent 2-sided estimate. While this makes estimates of the output gap behave in a more "orderly" fashion at the end of sample, the net effect on the accuracy of the estimated output gap is unclear.

One way to better understand the effects of these two changes at the end of samples is to compare the resulting 1-sided filter to the 1- and 2-sided HP filters examined in section 1.[20] As shown in Figure 10, these modifications cause the EMVF to put much less weight on the last few observations of the sample than the 1-sided HP filter, and overall make its weights more closely resemble those of the 2-sided HP filter. If we look in the frequency domain, however, we see that this change causes the 1-sided EMVF to pass more of the undesired low-frequency or "trend" components than either of the HP filters. In fact, Figure 11 shows that the squared-gain of the filter is greater than 0.2 for all frequencies.

---

[20] The remainder of this section expands the analysis of the EMVF filter properties presented in Butler (1996) to include the effects of the recursive updating restriction.

Figure 10

MA representation of the EMV & HP (λ = 1,600) filters (as a function of sample position)

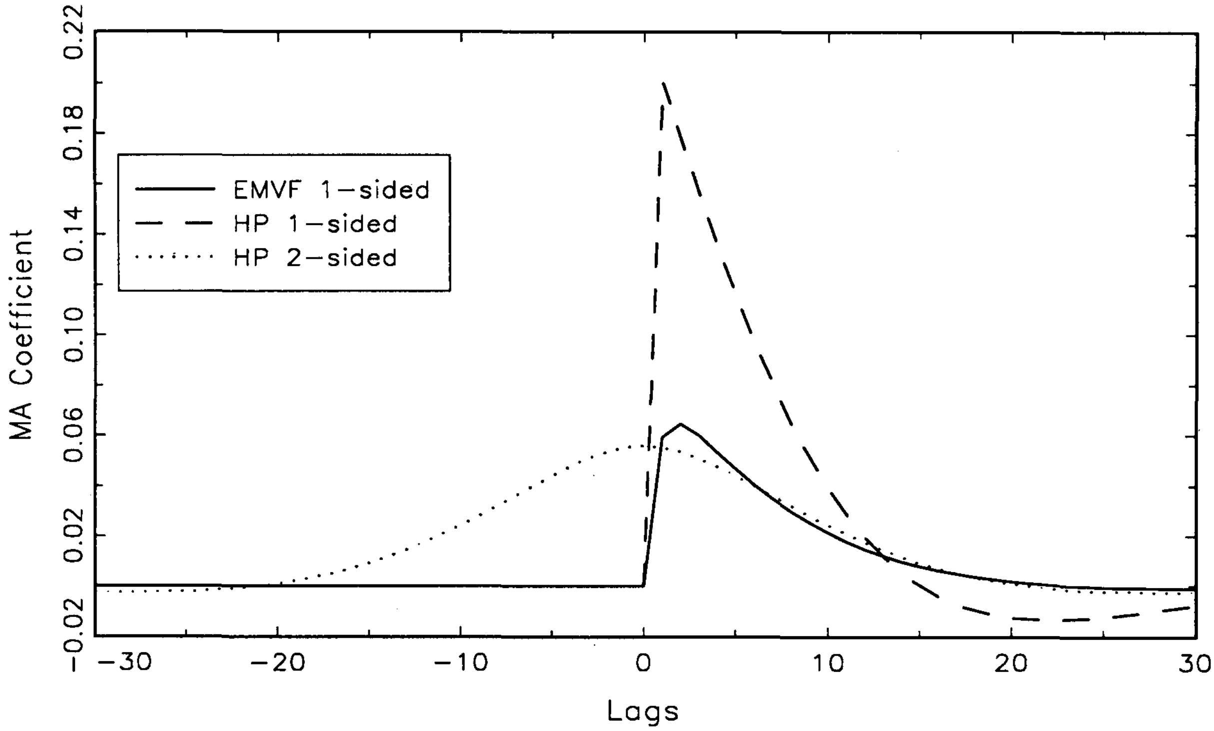


Figure 11

Squared gain of the EMV & HP (λ = 1,600) filters

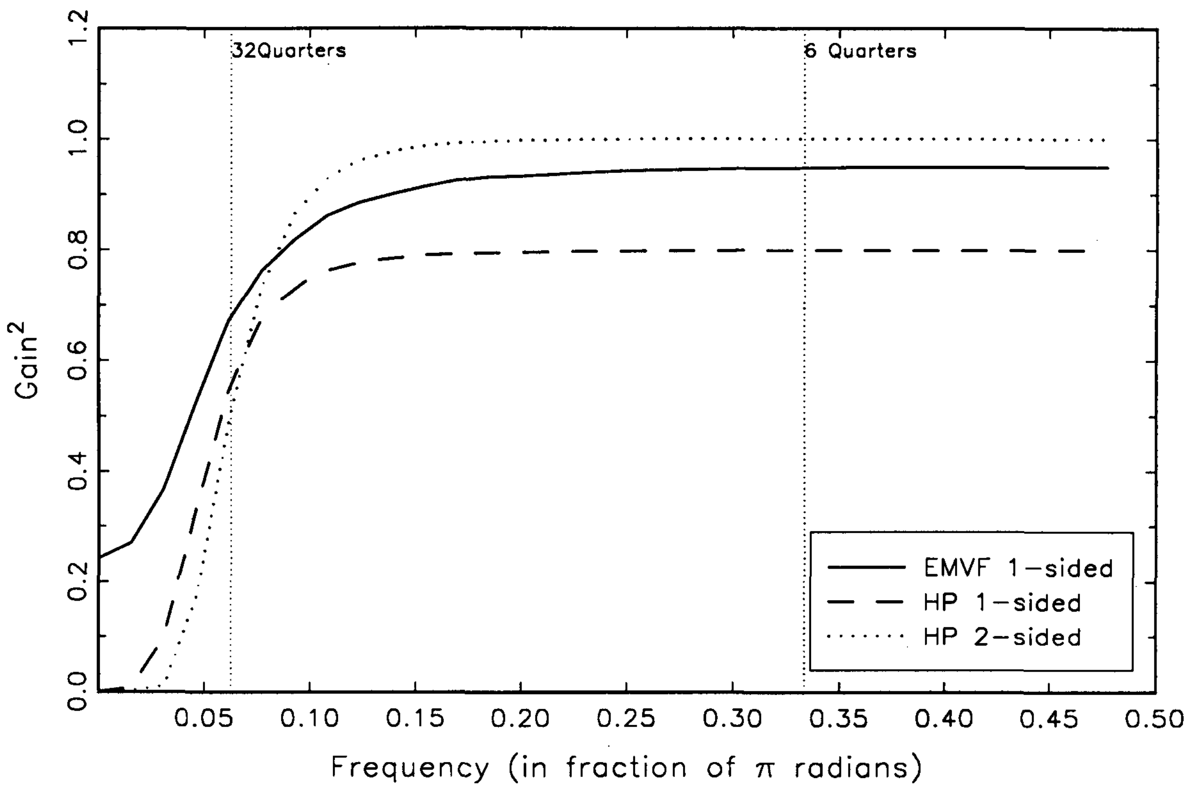


Figure 12

Spectrum of the EMVF gap and its components (Ar (3) fit for 1954Q4 - 1996Q4)

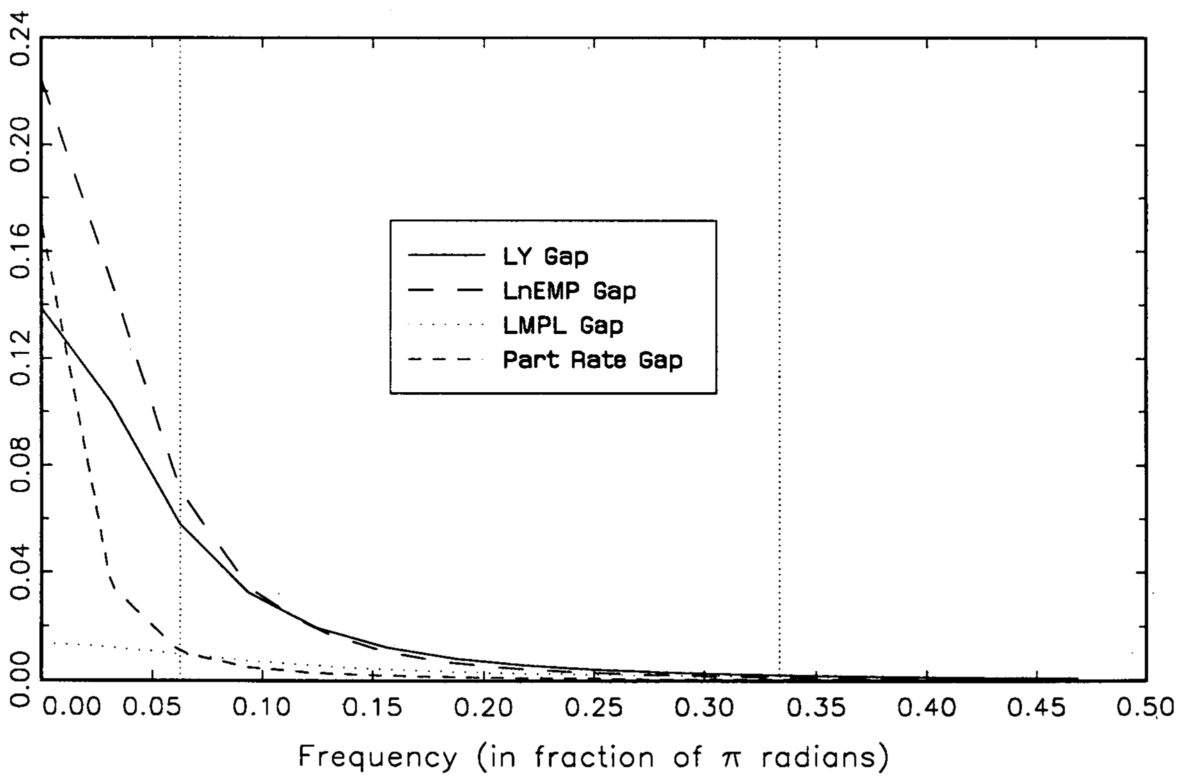
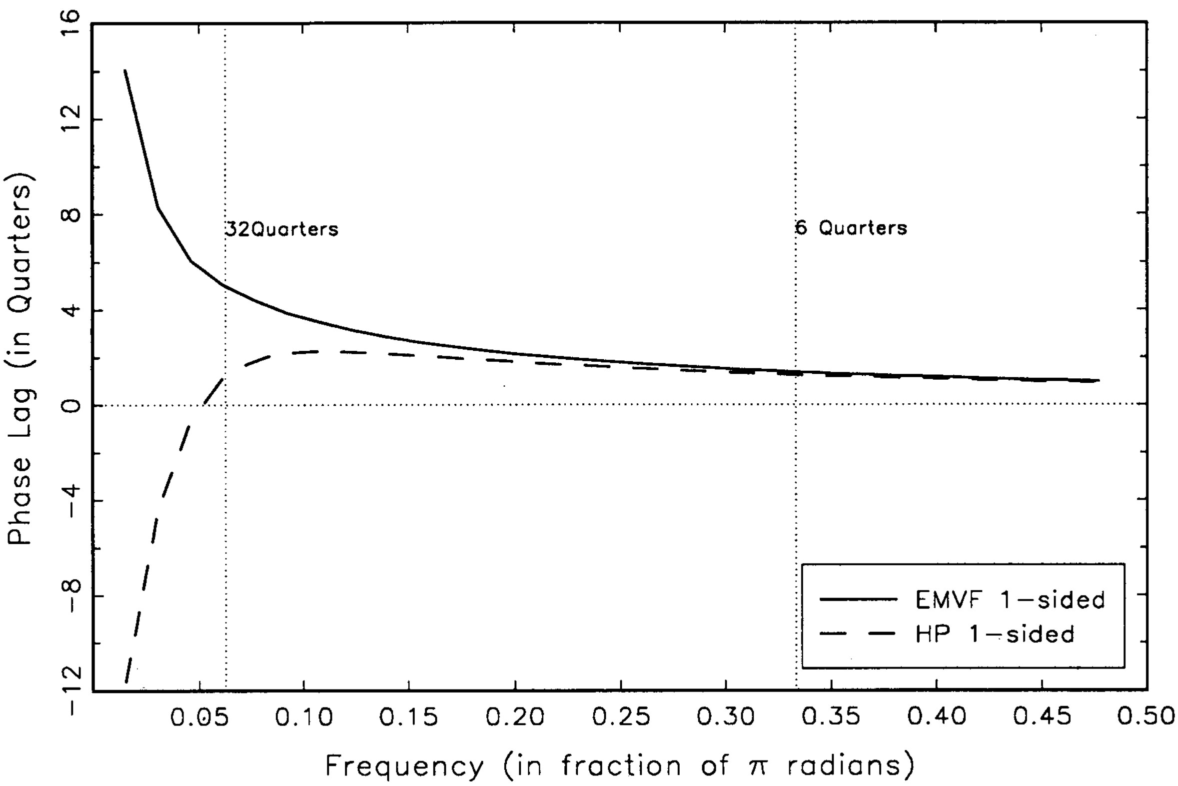


Figure 13

Phase shift of the EMV & HP (λ = 1,600) filters



The result of this is clearly shown in Figure 12, where we see that both the EMVF output gap and each of its three components appear to be dominated by low-frequency movements not normally associated with business cycles. Compared with Figure 7, it appears that the end-of-sample modifications of the EMVF make the filter much less able to isolate fluctuations at business-cycle frequencies than its simple HP filter counterpart. One way to quantify this effect is to use the estimated spectrum to calculate the correlation of the "measured" EMVF gap with that of an "ideally filtered" gap which perfectly isolated the business cycle frequencies. The result is that the measured EMVF output gap has a correlation with the "ideally filtered" gap of 31.4%, while its two filtered components have correlations of 24.9% (employment gap) and 44.0% (labour productivity gap).

Another implication of the differences in weights between the 1-sided HP and the 1-sided EMVF filter is that, since relatively more of its weight comes from observations with greater lags, the EMVF must have a greater phase shift than the HP filter at the end of sample. This in turn implies that the measured EMVF output gap will tend to lag the true output gap by more than the measured HP output gap. The extent of this difference depends on the frequency of the data series as shown in Figure 13. For all but the lowest of the business-cycle and the sub-business-cycle frequencies, the difference between the two is small and is roughly constant at a lag of about 2 quarters. For lower frequencies, however, the phase shift of HP falls to zero and then becomes negative, while that of the EMVF reaches 5 quarters by the lower bound of the business cycle frequencies and increases rapidly thereafter. If we weight these different phase shift by the relative importance of the different frequencies in measured output gaps, we obtain a weighted average measure of the overall phase lag for these different measures. This gives an overall phase lag of roughly 0 for the 1-sided HP filter, compared to a lag of 3.3 quarters for the EMVF output gap, 3.8 for the EMVF employment rate gap and 2.1 for the EMVF labour productivity gap.

2.4 TOFU

In Section 1, we looked at how well a time-series method (in this case, the HP filter) measured output gaps, and we concluded that by itself it did not produce very reliable estimates for policy makers. In Section 2 so far, we have looked at how adding sources of structural information to HP and related filters could improve the situation. One of the strengths of the MHPF approach is that it clearly states the problem which the resulting estimate of the output gap solves. However, some components of that optimization problem may be easier to accept than others. It seems reasonable to us that the estimated output gap be as consistent as possible with one or more structural economic relationships. The justification for the smoothing portion of the filter is more difficult, as was noted in Section 1, and it is unclear how helpful the particular assumptions of the HP filter are in identifying output gaps in the presence of structural information. Furthermore, as the complexity of the filter increases, the question of how to choose the parameters controlling the filter's behaviour becomes more difficult. While it is conceptually straightforward to estimate the filter parameters jointly with the structural relationships (as in Côté and Hostland (1994)) this can be quite difficult to implement. Accordingly, it would be useful to restate the optimization problem in a way that allows for easier estimation and has a filtering interpretation that is easier to justify.[21] This leads to an alternative to the HP filter which we refer to as TOFU.

From equation (13) we see that we can estimate the output gap in terms of the observable variable Δy_t if we can identify $\tilde{A}(L)$. Presumably, if we know of an economic relationship that involves the output gap, we could use this to define an optimal estimate of $\tilde{A}(L)$, call it $\hat{A}(L)$. For example, we might consider a Phillips curve of the form:

[21] This section draws heavily on van Norden (1995).

$$\pi_t = \alpha_0 + \gamma(q_t - y_t) + B(L).\pi_{t-1} + C(L).z_t + e_t \tag{29}$$

where z_t is a vector of additional observable variables, e_t is an i.i.d. mean zero error term and $B(L)$, $C(L)$ are one-sided polynomials in non-negative powers of L. We could substitute equation (13) into equation (29) to get:

$$\pi_t = \alpha_0 + \gamma(\tilde{A}(L).\Delta y_t) + B(L).\pi_{t-1} + C(L).z_t + \tilde{e}_t \tag{30}$$

where $\tilde{e}_t = e_t + \gamma\varepsilon_t$. Equation (30) can now be estimated by conventional methods to obtain optimal estimates of $\tilde{A}(L)$, since it is now specified entirely in terms of observable variables. This would allow us to estimate $\hat{A}(L).\Delta y_t$ and use equation (13) to estimate the output gap.[22]

We refer to this estimator of the output gap as TOFU; a Trivial Optimal Filter that may be Useful. It is optimal in the sense that estimation by maximum likelihood is straightforward, so our estimates of $\tilde{A}(L)$ will be efficient. The estimator imposes quite general assumptions on the time series properties of the series involved, so the restrictions are hopefully reasonable. It incorporates a simple structural relationship in order to identify the output gap. Furthermore, if we wished to estimate the output gap at the end-of-sample, we can simply replace $\tilde{A}(L)$ with $\overline{A}(L)$ (i.e. use only lagged values of Δy_t). This estimator therefore potentially avoids some of the problem mentioned at the outset of this section.

Of course, the TOFU estimate of the output gap is related to the estimate which would be obtained by "inverting" the structural relationships. The difference is simply that inversion calculates the implicit value of the output gap which would exactly fit the structural relationship. TOFU is therefore a half-way house between such methods and the MHPF methods of Section 2.1. The HP methods are optimal filters only for quite special cases whereas the TOFU methods gives us the optimal linear filter estimate of the output. TOFU offers less smoothing than MHPF methods but more than simple inversion of the structural equations.

Unfortunately, the information gained from structural relationships is contaminated by considerable "noise". Inversion of structural equations is therefore rarely used as a guide to policy, since the resulting estimates of the output gap are usually considered to be too volatile to be of practical use. Whether estimated TOFU filters can reduce this noise enough to be a useful tool for policymakers remains to be seen. If so, they may offer a tractable alternative to the MHPF methods. If not, it suggests that MHPF estimates may be dominated by the arbitrary assumptions they impose on the dynamics of the output gap rather than on the information coming from structural relationships. This suggests looking for other sources of information on these dynamics, which we turn to in the next section.

---

[22] We should note two minor caveats. First, equation (30) identifies $\tilde{A}(L).\Delta y_t$ only up to the scaling factor γ. Strictly speaking, therefore, we only recover an index of the output gap. This should be reasonable for the purpose of, say, deciding whether interest rates should be higher or lower to achieve a given target, since the current value of the index can be readily compared to its historical values. Second, consistent estimation of this relationship requires an implicit assumption. OLS estimation requires $\mathrm{cov}(e_t, \Delta y_{t-j}) = 0 \ \forall j$ for consistency. If this condition is not satisfied, then instruments for Δy will be required for estimation. Consistent IV estimation in turn will depend on the assumption that the chosen instruments are valid. This estimator can be extended in a number of ways. If the output gap were to enter the structural equation in a non-linear fashion, we could estimate the system via GMM rather than least-squares techniques. If we had a series of structural equations involving the output gap, we could estimate them simultaneously subject to cross-equation restriction on the coefficients of Δy_t.

3. Using long-run restrictions to estimate the output gap

In this section, we discuss approaches based on long-run restrictions imposed on a VAR. These approaches allow the identification of structural shocks and structural components on the basis of a limited number of economic restrictions imposed on an estimated VAR. The chosen restrictions can be ones that are widely agreed upon in the literature. No arbitrary mechanical filter has to be imposed on the data. Other characteristics distinguishing the methods discussed in this section from those based on mechanical filters are that they do not suffer from obvious end-of-sample problems and that they provide forecasted values of the output gap.

In Section 3.1, we discuss the method based on long-run restrictions imposed on output (LRRO) put forward by Blanchard and Quah (1989), Shapiro and Watson (1988), and King et al. (1991). This method is compared with two VAR-based alternatives: the multivariate Beveridge-Nelson method (MBN) and Cochrane's (1994) method (CO). We argue that one important advantage of the LRRO approach over the MBN and CO approaches is that it allows for the diffusion process of shocks to potential output to be estimated. Section 3.2 considers an application of the LRRO methodology to Canadian data. Many of the arguments used in Sections 3.1 and 3.2 are drawn from Dupasquier, Guay and St-Amant (1997).

In Section 3.3, we present a method involving restrictions imposed on real output and inflation (LRROI) giving an output gap corresponding to that part of the cyclical component of real output which is associated with the trend of inflation. This should be of interest for policymakers concerned about an output gap associated with movements in that trend as opposed to cyclical movements of output unrelated with that trend. This method is discussed in more detail by Lalonde, Page and St-Amant (forthcoming).

3.1 The LRRO, CO and MBN methodologies[23]

Let Z_t be a n x 1 stationary vector including a n_1-vector of I(1) variables and a n_2-vector of I(0) variables such that $Z_t = \left(\Delta X_{1t}', X_{2t}'\right)'$. By the Wold decomposition theorem, Z_t can be expressed as the following reduced form:

$$Z_t = \delta(t) + C(L)\varepsilon_t \tag{31}$$

where $\delta(t)$ is deterministic, $C(L) = \sum_{i=0}^{\infty} C_i L^i$ is a matrix of polynomial lags, $C_0 = I_n$ is the identity matrix, the vector ε_t is the one-step-ahead forecast errors in Z_t given information on lagged values of Z_t, $E(\varepsilon_t) = 0$ and $E\left(\varepsilon_t \varepsilon_t'\right) = \Omega$ with Ω positive definite. We suppose that the determinantal polynomial $|C(L)|$ has all its roots on or outside the unit circle, which rules out the non-fundamental representations emphasized by Lippi and Reichlin (1993).

Beveridge and Nelson (1981) show that equation (31) can be decomposed into a long-run component and a transitory component:

$$Z_t = \delta(t) + C(1)\varepsilon_t + C^*(L)\varepsilon_t \tag{32}$$

with $C(1) = \sum_{i=0}^{\infty} C_i$ and $C^*(L) = C(L) - C(1)$. We define $C_1(1)$ as the long-run multiplier of the

---

[23] See Cogley (1996) for another comparison of the MBN and CO methodologies.

vector X_{1t}. If the rank of $C_1(1)$ is less than n_1, there exists at least one linear combination of the elements in X_{1t} that is I(0). In other words, there exists at least one cointegration relationship between these variables.

The LRRO approach assumes that Z_t has the following structural representation:

$$Z_t = \delta(t) + \Gamma(L)\eta_t \tag{33}$$

with η_t an n-vector of structural shocks, $E(\eta_t) = 0$, and $E\left(\eta_t \eta_t'\right) = I_n$ (a simple normalization). We can retrieve the structural form (equation (33)) from the estimated reduced form by using the following relationships: $\Gamma_0\Gamma_0' = \Omega$, $\varepsilon_t = \Gamma_0\eta_t$, and $C(L) = \Gamma(L)\Gamma_0^{-1}$.

The long-run covariance matrix of the reduced form is equal to $C(1)\Omega C(1)'$. From equations (31) and (33) we have:

$$C(1)\Omega C(1)' = \Gamma(1)\Gamma(1)' \tag{34}$$

This relation suggests that we can identify the matrix Γ_0 with an appropriate number of restrictions on the long-run covariance matrix of the structural form. Blanchard and Quah (1989) and Shapiro and Watson (1988) use long-run restrictions to identify shocks with $C(1)$ having full rank. King et al. (1991) work in a context where the rank of $C(1)$ is less then n_1 and use cointegration restrictions.

Let us assume that the log of real output is the first variable in the vector Z_t. It is then equal to

$$\Delta y_t = \mu_y + \Gamma_1^p(L)\eta_t^p + \Gamma_1^c(L)\eta_t^c \tag{35}$$

where η_t^p is the vector of permanent shocks affecting output, η_t^c is the vector of shocks having only transitory effects on output, and $\left\{\Gamma_1^p(L), \Gamma_1^p(L)\right\}$ reflect the dynamic effects of these shocks. Potential output growth based on the LRRO method can then be defined as:

$$\Delta y_t^p = \mu_y + \Gamma_1^p(L)\eta_t^p \tag{36}$$

Thus, "potential output" corresponds to the permanent component of output. The part of output due to purely transitory shocks is defined as the "output gap."

The MBN decomposition defines potential output as the level of real output that is reached after all transitory dynamics have worked themselves out. With reference to equation (32), where real output is the first element of Z_t, we write the following decomposition: $\Delta y_t = \mu_y + C_1(1)\varepsilon_t + C_1^*(L)\varepsilon_t$ (37)

Potential output can be defined as the first two terms on the right-hand side of equation (37):

$$\Delta y_t^p = \mu_y + C_1(1)\varepsilon_t \qquad (38)$$

It is thus simply a random walk with drift.

Note that the MBN approach gives an output gap that is sensitive to the choice of variables included in the VAR. In general, the more information that is brought into the VAR, the more important the transitory component will be. This is not the case with the LRRO approach. Adding additional information may or may not add to the importance of the cyclical component.

Cochrane (1994) uses a two-variable VAR including GNP and consumption to identify the permanent and transitory components of GNP. The bivariate representation is augmented with lags of the ratio of consumption to GNP. The permanent income theory implies that consumption is a random walk (for a constant real interest rate). In addition, if we assume that GNP and consumption are cointegrated, then fluctuations in GNP with consumption unchanged must be perceived as transitory. It is on that basis that Cochrane decomposes real GNP into permanent and transitory components. To extract potential output, the errors of the VAR are orthogonalized so that consumption does not respond contemporaneously to GNP shocks.

Cochrane shows that, if GNP and consumption are cointegrated and consumption is a random walk, identification based on the LRRO method and conventional orthogonalization (i.e. a Choleski decomposition) amount to the same thing. Moreover, if consumption is a pure random walk, Cochrane's decomposition corresponds exactly to the Beveridge-Nelson decomposition based on output and consumption.

In order to better compare the LRRO approach with the CO and MBN approaches, let us first write the structural form (equation (33)) in terms of the log of real GDP (y_t) and the log of real consumption (c_t) decomposed between permanent and transitory shocks (we assume that y_t and c_t are cointegrated):

$$\Delta y_t = \mu_y + \Gamma_y^p(1)\eta_t^p + \Gamma_y^{p*}(L)\eta_t^p + \Gamma_y^c(L)\eta_t^c \qquad (39)$$

$$\Delta c_t = \mu_c + \Gamma_c^p(1)\eta_t^p + \Gamma_c^{p*}(L)\eta_t^p + \Gamma_c^c(L)\eta_t^c \qquad (40)$$

where $\Gamma^p(1)$ is the long-run multiplier of permanent shocks and $\Gamma_c^{p*}(L) = \Gamma^p(L) - \Gamma^p(1)$ is their transitory component. The MBN method considers only the first component of the permanent shocks plus the drift term, i.e. $\mu + \Gamma_y^p(1)\eta_t^p$. The LRRO approach is different in that it also includes the dynamics of permanent shocks to real output $\left(\Gamma_y^{p*}(L)\right)$ in potential output.

With the CO approach, potential output is constrained to be a random walk to the extent that consumption is a random walk. Indeed, the validity of the permanent income hypothesis would imply that the last two terms of equation (39) are equal to zero and that $\Gamma_y^p(1) = \Gamma_c^p(1)$. It is not clear what the CO decomposition would correspond to if consumption is not a random walk.[24]

As pointed out by Lippi and Reichlin (1994), modelling the trend in real output as a random walk is inconsistent with most economists' interpretation of productivity growth. Indeed, it is

---

[24] Stochastic growth models – such as in King et al. (1988) or King et al. (1991) – imply that the ratio of the log of GNP to the log of consumption is stationary but that consumption is not a random walk because the real interest rate is not constant. In these models, the transitory component of permanent shocks to consumption is not equal to zero. The LRRO decomposition is compatible with the predictions of these models.

generally believed that technology shocks are absorbed gradually by the economy. Adjustment costs for capital and labour, learning and diffusion processes, habit formation, and time to build imply richer dynamics than a random walk for these shocks. Again, a decisive advantage of the LRRO approach is that it lets the data determine the shape of the diffusion process of permanent shocks.[25]

One implication of defining potential output as a random walk with drift is that, when the contemporary effect of a positive permanent shock is smaller (greater) than its long-run effect, the output gap, defined as observed output minus potential, is negative (positive). For example, a positive technological shock whose short-term impact is smaller than its long-term impact will cause a transitory negative output gap. Many researchers and/or policymakers will find that this feature of the MBN and CO approaches (in the latter case under the assumption that consumption is a random walk) reduces their attractiveness. It will often appear preferable to include the diffusion process associated with permanent shocks in potential output since the economy is likely to remain on its production possibility frontier as adjustments unfold. There should be no reason for the trend of inflation to change during that adjustment process.

3.2 An application of the LRRO approach to Canadian data

For our applications of the LRRO methodology to Canadian data, we assume that the growth rate of real output (Δy_t) follows a stationary stochastic process responding to two types of structural shocks: permanent (ε_p) and transitory (ε_T). Also included in the estimated VARs are the first differences of, respectively, inflation ($\Delta\pi$), the unemployment rate (Δu) and the real interest rate (Δr). We assume that these series are I(0) and that there is no cointegration involved.[26]

Adding money or the exchange rate to the estimated VARs would have little impact on the results. In selecting the variables, one hopes to include in the VAR the information necessary to identify the structural components of interest. Of course, there is a cost in adding information in terms of lost degrees of freedom and less precise estimates.

It is important to note that the assumptions made on the level of integration of the series could be changed and that they are not part of the LRRO methodology per se (except for real output which has to be I(1)).[27] Changing these assumptions could have a significant impact on the results.

Our objective is to illustrate the methodology. A practitioner interested in using the approach to estimate the output gap might well choose to use additional information and be able to obtain a better estimate of the output gap.

The structural shocks and the variables used in the VAR can be expressed in the following vector form:

$$\eta_t = \begin{bmatrix} \varepsilon_P \\ \varepsilon_{T_1} \\ \varepsilon_{T_2} \\ \varepsilon_{T_3} \end{bmatrix} \text{ and } Z_t = \begin{bmatrix} \Delta y \\ \Delta\pi \\ \Delta u \\ \Delta r \end{bmatrix} \quad (41)$$

---

25 Kuttner (1994) proposes a method based on the univariate unobserved stochastic trend decomposition of Watson (1986) augmented with a Phillips curve equation. As with the Beveridge-Nelson decomposition, Kuttner's approach constrains potential output to follow a random walk process.

26 Unit root tests support these assumptions. Results are available on request.

27 Indeed, DeSerres, Guay and St-Amant (1995) assume that money growth is I(0).

We use quarterly data on real GDP. Our measure of inflation is the total CPI. The real interest rate is proxied by the overnight rate (see Armour et al. (1996) for a discussion of that series) minus inflation (quarterly growth rate). Our sample extends from the first quarter of 1970 to the fourth quarter of 1996 so that we focus on the flexible exchange rate period. The estimation of the output gap based on the LRRO methodology is fairly robust to the choice of the sample period, however.

The autoregressive reduced-form of the model is first estimated:

$$Z_t = \sum_{i=1}^{q} \Pi_i Z_{t-i} + e_t \tag{42}$$

with q the number of lags and e_t a vector of estimated residuals with $E(e_t e_t) = \Sigma$.

It is crucial that the estimated VARs include enough lags. Indeed, Monte Carlo simulations by DeSerres and Guay (1995) show that using a too parsimonious lag structure can significantly bias the estimation of the structural components. We decided to use 8 lags. However, we verified that using 6 or 10 lags had little impact on the results.

The LRRO approach involves the identification of structural shocks (ε_t) from reduced-form shocks (e_t) and their variance. For this, we need to provide enough identifying restrictions to evaluate the 16 elements in Γ_0. Given that Σ is symmetric, we need to impose 6 additional restrictions. The matrix of long-run effects of reduced-form shocks, $C(1)$, is related to the equivalent matrix of structural shocks, $\Gamma(1)$, as follows:

$$\Gamma(1) = C(1)\Gamma_0 \tag{43}$$

where the matrix $C(1)$ is calculated from the estimated VAR. To identify the system we simply impose the condition that $\Gamma(1)$ is triangular, i.e. three shocks have no long-run effect on real output and two have no long-run effect on inflation. There are then three transitory components of real output which we do not need to identify separately.

Figure 14 presents the impulse response of real output to a one standard deviation permanent output shock. (Blanchard and Quah suggest interpreting this as an aggregate supply shock.) The horizontal axis represents the number of years. Confidence intervals were generated using Monte Carlo simulations in RATS with 1,000 replications.

The important message of Figure 14 is that permanent shocks are characterized by statistically significant dynamics, i.e., potential output has richer dynamics than a simple random walk.[28] As mentioned above, this could be due to factors such as adjustment costs on capital and labour, learning, habit formation, or time to build. One implication of the rejection of the random walk assumption is that methods which do not take the diffusion process of permanent shocks into account could miss an important part of potential output. Indeed, Dupasquier, Guay and St-Amant (1997) show that the correlation between the output gaps calculated on the basis of the LRRO, CO and MBN approaches applied to US data is relatively small. Part of this can be attributed to the different treatments of the diffusion process of permanent shocks

Figure 15 shows the output gaps calculated on the basis of the LRRO methodology as it is applied in this paper, together with 90 and 67% confidence intervals.

[28] Dupasquier, Guay and St-Amant (1997), Blanchard and Quah (1989) and Gali (1993), among others, report similar results for the US economy.

Figure 14

Response of real GDP to a permanent output shock

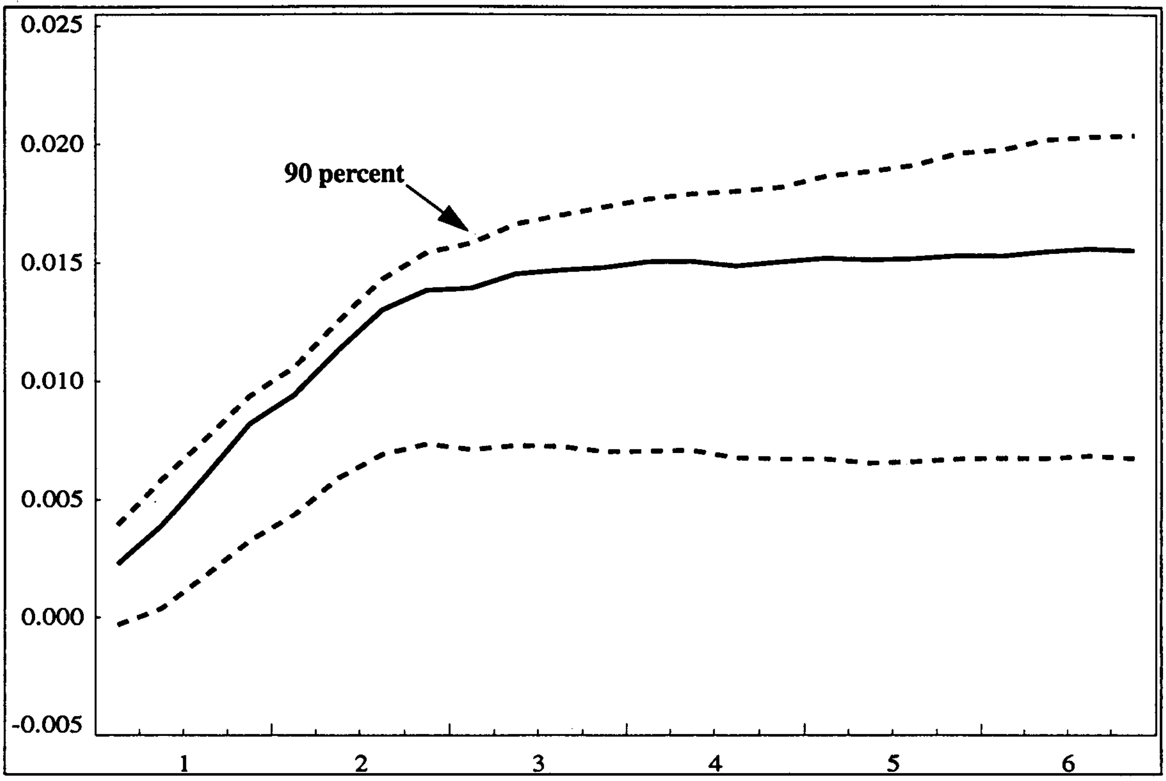
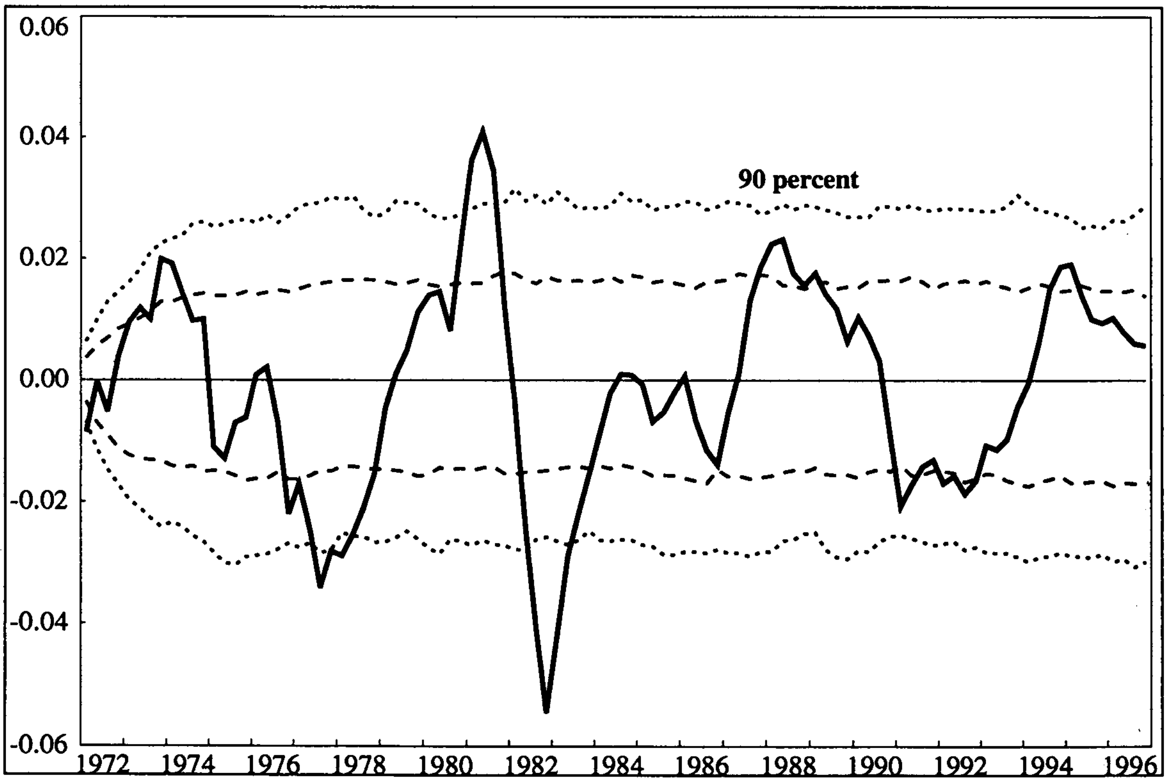


Figure 15

LRRO-based Canadian output gap



One important message of Figure 15 is that there is a high degree of uncertainty surrounding the estimation of the output gap. Dupasquier, Guay and St-Amant (1997) reach the same conclusion using US data. Staiger, Stock and Watson (1996), using a different methodology, arrive at a similar conclusion concerning the estimation of the NAIRU for the United States. Some of that uncertainty is attributable to the large number of lags that have to be included in the estimated VARs. As mentioned above, DeSerres and Guay (1995) show that many lags have to be used to provide an unbiased decomposition into permanent and transitory components with structural VARs. To a large extent, the purpose of these lags is to approximate the moving-average part of the underlying DGP. Preliminary results obtained at the Bank of Canada suggest that the estimation of VARMAs instead of VARs could reduce parameter uncertainty by allowing the use of more parsimonious models.

Still, there are episodes of significant output gaps at either the 90% level or the 67% level. These output gaps appear reasonable in that positive output gaps are associated with episodes of accelerating inflation while negative output gaps correspond to episodes of decelerating inflation. Interestingly, we find that the Canadian economy was either at potential or in slight excess demand in the mid-1990s.

It should be of interest to see whether the structural shocks used to calculate the output gap account for an important part of the variance of inflation at different horizons. That information is presented in Table 2.

Table 2

Variance decomposition of Canada's inflation (LRRO method)

Relative contribution of the different types of shocks, in percentage points

Horizon (quarters)	Permanent output shock	Transitory output shock
1	11	89
	(0-43)*	(57-100)
4	8	92
	(2-36)	(64-98)
8	6	94
	(4-30)	(69-97)
16	13	87
	(5-49)	(51-95)
32	23	77
	(5-66)	34-95)
long-term	37	63
	(3-84)	(16-97)

* 90% confidence interval.

We can see that transitory shocks affecting real output account for a large part of the variance of Canadian inflation. This result suggests that the component of output that we identify includes much of the information one would want to include in the output gap. Note however that there appears to be some additional information related to permanent shocks, especially at longer horizons. Lalonde, Page and St-Amant (forthcoming) discuss this and propose another approach which does not impose that the output gap is part of the cyclical component of output.

Instead, in this paper we keep our focus on methods imposing that the output gap is stationary. Although Table 2 suggests that the output gap based on the LRRO approach accounts for a large fraction of fluctuations in the trend of inflation, it may also be that a part of that gap is unrelated to the trend. For example, it might include very-high-frequency cycles having little to do with the trend. This leads us to another method which involves imposing restrictions on both real output *and* inflation in order to produce an output gap that is constrained to be associated with movements in the trend of inflation.

3.3 The LRROI approach

It is possible, with the four-variables system presented in Section 3.2, to identify shocks that have no long-run effects on real output but affect the trend of inflation. Real output was indeed decomposed in the following way:

$$\Delta y_t = \mu_y + \Gamma_y^p(1)\eta_t^p + \Gamma_y^{p*}(L)\eta_t^p + \Gamma_y^{cp}(L)\eta_t^{cp} + \Gamma_y^{cc}(L)\eta_t^{cc} \tag{44}$$

with certain shocks having no long-run effect on real output but changing the trend of inflation (η_t^{cp}), certain shocks having no long-run effect on real output *and* inflation (η_t^{cc}),[29] and certain shocks having long-run effects on real output but whose effect on inflation is left unconstrained (η_t^p). The component $\Gamma_y^{cp}(L)\eta_t^{cp}$ could be used as a measure of the output gap. This is what we call the LRROI method. In Section 3.2 we were not interested in distinguishing $\Gamma_y^{cp}(L)\eta_t^{cp}$ from $\Gamma_y^{cc}(L)\eta_t^{cc}$ so that we simply added these components to form one component that was the LRRO output gap.

The LRROI method gives a measure of the output gap that is more constrained than the LRRO method in that it combines restrictions on real output *and* inflation. While the LRRO approach only requires that real output is I(1), the additional assumption that inflation is better characterized as being I(1) over the sample is a necessary feature of the LRROI approach. Of course, that assumption is not uncontroversial. However, we think it is reasonable. Assuming that inflation is I(0) is in some sense equivalent to assuming that, whatever the actions of the monetary authorities, it has to go back to a constant mean. We do not think this is realistic. At least, inflation has to be modelled as a process, the mean of which changes through time. One way to do it is to assume that it is I(1).

The LRROI method proposes a measure of the output gap that should be attractive for policymakers who focus on that part of the cyclical component of real output that is associated with movements in the trend of inflation as opposed to short-run fluctuations of that series. Instead, the LRRO method provides a measure of the output gap including cyclical movements that may not be of interest to such policymakers. In some sense, the LRROI method could provide policymakers with a less "noisy" indicator of changes in the trend of inflation. Short-run fluctuations of inflation could be caused by factors such as transitory fluctuations in the exchange rate or changes of indirect taxes which may not require any change in the policy stance. A caveat to this is of course that there might be instances when policymakers will be interested in reacting to fluctuations of output associated with transitory changes of inflation. Unlike the LRRO gap, the LRROI gap would not include the effect of such shocks.

Figure 16 presents the LRROI output gap calculated on the basis of the Canadian data used in Section 3.3. Figure 17 compares that output gap with the one resulting from the LRRO approach to the gap from the one-sided HP filter. Table 3 presents the variance decomposition of inflation corresponding to the LRROI decomposition.

We can see in Figure 16 that there is as much uncertainty surrounding the estimation of the LRROI output gap as there is with the LRRO gap. We also see that the LRROI gaps are generally smaller than those produced by the LRRO method. This indicates that some cycles included in the LRRO gap are not related to the trend of inflation. Finally, the variance decomposition of inflation suggests that the LRROI approach includes most of the information related to the cyclical component of output that is relevant for monitoring or forecasting medium-and-long-term inflation. However, the

---

[29] Since there are two shocks which have no long-run impact on the trend of output or inflation (the last two of equation (44)), we simply consider their sum.

Figure 16

LRROI-based Canadian output gap

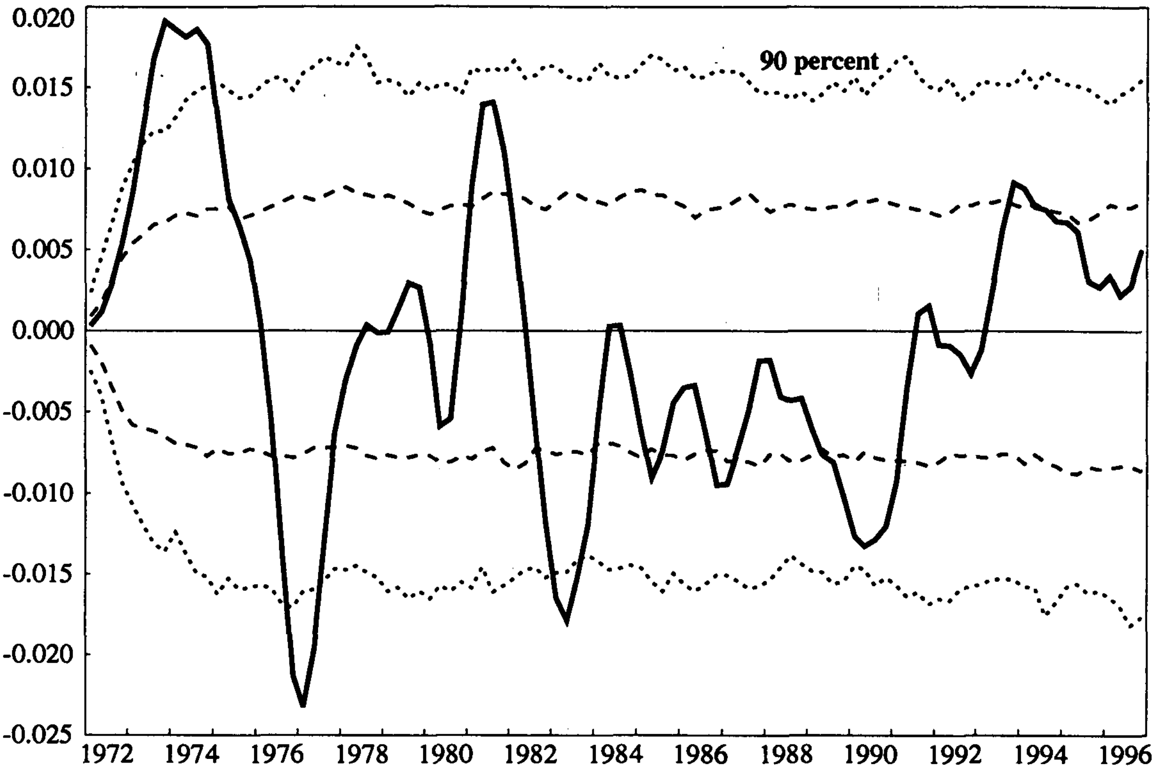


Figure 17

LRRO, LRROI and one-sided HP-based Canadian output gaps

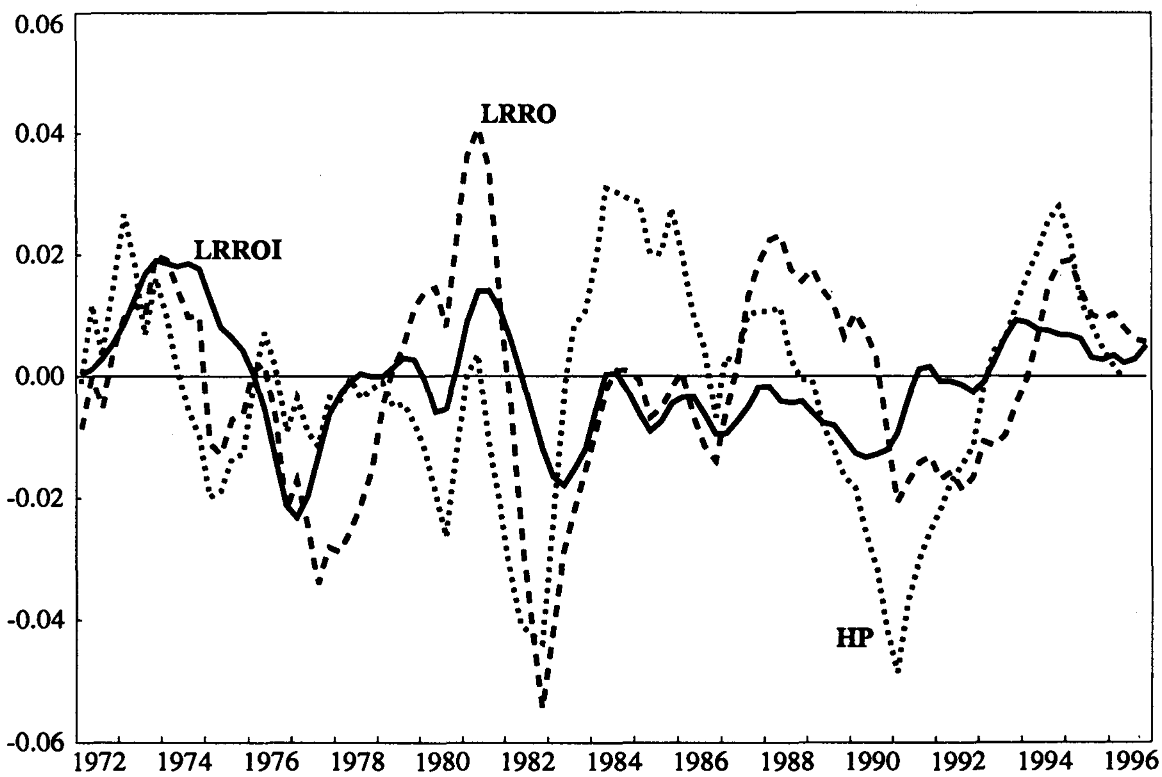


Table 3

Variance decomposition of Canada's inflation (LRROI method)
Relative contribution of the different types of shocks, in percentage points

Horizon (quarters)	Permanent output shock	Transitory output shocks affecting the trend of inflation	Transitory output shocks that do not affect the trend of inflation
1	11 (0-43)*	54 (6-88)	35 (3-81)
4	8 (2-36)	63 (18-85)	29 (7-68)
8	6 (4-30)	70 (33-80)	24 (11-53)
16	13 (5-49)	72 (34-80)	15 (8-35)
32	23 (5-66)	68 (27-83)	9 (4-23)
long-term	37 (3-84)	61 (16-95)	1 (0-4)

* 90% confidence interval.

LRRO gap appears to include non-negligible information about movements of inflation lasting as much as a couple of years.

To conclude this section we would like to emphasize the fact that the LRRO and LRROI methods are two variants of a more general VAR-based approach. Other variants could be considered. For example, one might want to take into account possible non-linearities in the relationship between output and inflation (e.g. Fillion and Léonard (1997)). Evans (1992) proposes a time-varying method allowing for such non-linearities. One might also want to include different sets of restrictions, including cointegration and bayesian types of restrictions, to ensure that the estimated gap is compatible with some specific macroeconomic model.

Summary

Much research has gone into measuring output gaps and much still remains to be done. On the basis of the results discussed, we think that there are three main lessons to be learned.

1. Univariate time-series methods, such as the HP filter, are not a reliable way to measure the output gap

Univariate methods rely on an arbitrary decomposition of a series into a trend and a cyclical component. However, changing the decomposition method can make large differences to the measured output gap, and economic theory usually has little or nothing to say about which method should be favoured. In addition, the causal relationship between potential and the gap can limit the information we can hope to gain about the current gap, and some popular economic models imply that univariate filters will never be able to give much information.

In the particular case of the HP filter, although this filter is thought to be close to an ideal high-pass filter, it does not accurately measure the components from business-cycle frequencies when series have the typical Granger shape. Although the HP filter can also be justified as an optimal filter for particular cases, these cases do not appear to be realistic approximations of output and the filter is generally not a reliable way to estimate the "cyclical" component. HP filters also behave very differently at the end of the sample, which is what policymakers care most about, and little is known about the trade-off between phase shift and smoothness at the end of sample.

2. Existing hybrid methods that combine univariate dynamic methods and structural relationships are not a panacea

In practical terms, existing hybrid methods have proved to be hard to estimate. Moreover, they may not be robust to alternative reasonable calibrations and it is difficult to calculate their appropriate confidence intervals. In the case of the EMVF, which is the filter used to estimate the output gap in the Bank of Canada's staff projections, its ability to isolate business cycle frequencies is worse than that of the HP filter and its estimates lag the true output gap by just under a year. These problems seem to be the result of features which were introduced to improve the filter's estimate of current and recent output gaps.

More generally, the "hybrid" approach is not only driven by a desire to include structural relationships, but also by a pragmatic desire for "smooth" estimates of the output gap. However, if the structural relationships are very informative then the "smoothing" assumptions may be unnecessary. Otherwise, it is hard to argue that these assumptions will be innocuous, for the reasons we mentioned above in the context of univariate methods. For that reason, and because they can incorporate the same sources of structural information, further investigation of TOFU methods may provide a good benchmark for hybrid methods. If TOFU estimates of the output gap are useful, then the strict (or ad hoc) filters used in existing hybrid methods are not required. If TOFU methods are not useful, then structural information alone is not sufficient to identify the output gap. Relying instead on ad hoc dynamic assumptions for identification raises questions about the reliability and significance of the resulting estimated output gaps.

3. Methods combining estimated dynamics with structural information offer an interesting alternative which needs to be further explored

Examples of these methods explored in this paper are VARs with long-run restrictions suggested by economic theory. Their advantages include an absence of arbitrary dynamic assumptions, straightforward estimation, and an ability to estimate both current and expected future output gaps. On the other hand, it is not always clear which variables have to be included in the VAR and work to date suggests that the estimated output gaps have wide confidence intervals comparable to those of other methods. More work is needed to evaluate the extension of these methods to VARMA models. It might also be of interest to explore VARs with time-varying parameters (as is proposed by Evans (1992)) in order to deal with possible non-linearities in structural or dynamic relationships.

The VAR-based methods considered in this paper impose relatively little economic structure on the data and allow the dynamics properties of the estimated output gaps to be data-determined. This is an advantage insofar as economic theory provides little guidance on what the dynamics should be. Of course, one has to assume that the estimated VAR includes the information relevant for the identification of the output gap.

On the other hand, there might be instances when practitioners would want to impose more economic structure on their estimation of the output gap by incorporating more economic relationships. To some extent, this could be accommodated in the VAR (or VARMA or VECM) framework. Many types of economic or statistical restrictions have been discussed in the literature including bayesian priors and cointegration relationships. Such restrictions could be used to ensure that the estimated output gap is broadly compatible with some structural model of interest.

However, the TOFU approach might provide a better framework for those interested in imposing detailed economic relationships on the data (such as a specific Phillips curve or a NAIRU equation). One advantage of using detailed economic structure is that the derived measures of potential output can then be embedded in a model that is consistent with that structure. As noted briefly in van Norden (1995), extensions of the TOFU approach to multiple detailed structural relationships appears to be straightforward, although this remains to be explored in detail. It should be kept in mind, however, that the usefulness of TOFU output gaps would depend on the validity of the economic structure imposed on the data and that economics provides very few non controversial structural relationships.

In our view, the VAR-based and TOFU approaches deserve further research, while our findings suggest that one should avoid univariate methods and move away from arbitrary smoothing methods wherever practicable. Variants of the VAR-based and TOFU approaches could be considered and their usefulness in terms of monitoring and projecting inflation evaluated.

Bibliography

Baxter, M. and R. G. King (1995): "Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series". *Working Paper*, No. 5022. National Bureau of Economic Research.

Beveridge, S. and C. R. Nelson (1981): "A New Approach to the Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle". *Journal of Monetary Economics*, No. 7, pp. 151-74.

Blanchard, O. J. and S. Fischer (1989): *Lectures on Macroeconomics*, MIT Press, Cambridge (Massachussetts).

Blanchard, O. J. and D. Quah (1989): "The Dynamic Effect of Aggregate Demand and Supply Disturbances". *American Economic Review*, No. 79, pp. 655-73.

Boschen, J. and L. Mills (1990): "Monetary Policy with a New View of Potential GNP". *Federal Reserve Bank of Philadelphia Business Review*, pp. 3-10, July-August.

Burns, A. M. and W. C. Mitchell (1946): *Measuring Business-Cycles*. New York: National Bureau of Economic Research.

Butler, L. (1996): "A Semi-Structural Method to Estimate Potential Output: Combining Economic Theory with a Time-Series Filter". *Technical Report*, No. 77, Bank of Canada.

Cochrane, J. H. (1994): "Permanent and Transitory Components of GNP and Stock Prices". *Quarterly Journal of Economics*, No. 61, pp. 241-65.

Cogley, T. (1996): "Evaluating Non-Structural Measures of the Business Cycle". Draft Paper, Federal Reserve Bank of San Francisco.

Cogley, T. and J. Nason (1995): "Effects of the Hodrick-Prescott filter on Trend and Difference Stationary Time Series: Implications for Business Cycle Research". *Journal of Economic Dynamics and Control*, No. 19, pp. 253-78.

Côté, D. and D. Hostland (1994): "Measuring potential output and the NAIRU as unobserved variables in systems framework", in *Economic Behaviour and Policy Choice Under Price Stability*, Bank of Canada, pp. 357-418.

DeSerres, A. and A. Guay (1995): "The Selection of the Truncation Lag in Structural VARs (or VECMs) with Long-Run Restrictions". *Working Paper*, No. 95-3, Bank of Canada.

DeSerres, A., A. Guay and P. St-Amant (1995): "Estimating and Projecting Potential Output using Structural VAR Methodology: the Case of the Mexican Economy". *Working Paper*, No. 95-2, Bank of Canada.

Dupasquier C., A. Guay and P. St-Amant (1997): "A Comparison of Alternative Methodologies for Estimating Potential Output and the Output Gap". *Working Paper*, No. 97-5, Bank of Canada.

Dupasquier. C., R. Lalonde and P. St-Amant (1997): "Les zones monétaires optimales: une application au Canada et aux États-Unis". Forthcoming.

Engle, R. F. and C. W. J. Granger (1987): "Cointegration and Error Correction: Representation, Estimation and Testing". *Econometrica*, No. 55 (March), pp. 251-76.

Evans, G. and L. Reichlin (1994): "Information, Forecasts, and Measurement of the Business Cycle". *Journal of Monetary Economics*, No. 33, pp. 233-54.

Evans, M. D. D. (1992): "The Changing Nature of the Output-Inflation Trade-off". NYU Stern School of Business, *Working Paper*, No. 92-17.

Gali, J. (1992): "How well does the IS-LM Model fit Postwar US Data?" *Quarterly Journal of Economics*, No. 107, pp. 709-38.

Giorno, C., P. Richardson, D. Reveare and Paul van den Noord (1995): "Estimating Potential Output, Output Gaps and Structural Budget Balances". OCED Economic Department, *Working Paper*, No. 152.

Gouriéroux, C. and M. Monfort (1995): "Séries Temporelles et Modèles Dynamiques". *Economica.*

Guay, A. and P. St-Amant (1996): "Do Mechanical Filters Provide a Good Approximation of Business Cycles?" Bank of Canada, *Technical Report*, No. 78.

Hamilton, J. D. (1994): *Time Series Analysis*. Princeton University Press.

Harvey, A. C. and A. Jaeger (1993): "Detrending, Stylized Facts and the Business Cycle". *Journal of Applied Econometrics*, No. 8, pp. 231-47.

Henderson, R. (1924): "A New Method of Graduation". *Actuarial Society of America Transactions*, No. 25, pp. 29-40.

Hendry, S. (1995): "Long-Run Demand for M1". *Working Paper*, No. 95-11, Bank of Canada.

Hodrick, R. J. and E. Prescott (1997): "Post-war U. S. Business-Cycles: An Empirical Investigation". *Journal of Money Credit and Banking*, forthcoming.

Jeong, J. and G. S. Maddala (1993): "A Perspective on Application of Bootstrap Methods in Econometrics", in G. S. Maddala, C. R. Rao and H. D. Vinod (eds.), *Handbook of Statistics*, Vol. 11, pp. 573-610, New York: North-Holland.

King, R. G., C. I. Plosser and S. T. Rebelo (1988): "Production Growth and Business Cycles: II. New Directions". *Journal of Monetary Economics*, No. 21(2-3), pp. 309-41.

King, R. G., G. H. Plosser, J. H. Stock and M. W. Watson (1991): "Stochastic Trends and Economic Fluctuations". *American Economic Review*, No. 81 (September), pp. 819-40.

King, R. G. and S. Rebelo (1993): "Low Frequency Filtering and Real Business-Cycles". *Journal of Economic Dynamics and Control*, No. 17, pp. 207-31.

Kuttner, K. N. (1994): "Estimating Potential Output as a Latent Variable". *Journal of Business and Economic Statistics*, No.12, pp. 361-68.

Lalonde, R., J. Page and P. St-Amant (1997): "Trois méthodes d'estimation de l'écart de production reposant sur des restrictions de long terme". Forthcoming.

Laxton, D. and R. Tetlow (1992): "A Simple Multivariate Filter for the Measurement of Potential Output". *Technical Report*, No. 59. Ottawa: Bank of Canada.

Laxton, D., D. Rose and R. Tetlow (1993):"Is the Canadian Phillips Curve non-linear?" Bank of Canada, *Working Paper*, No. 93-7.

Lippi, M. and L. Reichlin (1993): "The Dynamic Effects of Aggregate Demand and Supply Disturbances: Comment". *American Economic Review*, No. 83, pp. 644-52.

Lippi, M. and Reichlin (1994): "Diffusion of Technical Change and the Decomposition of Output into Trend and Cycle". *Review of Economic Studies*, No. 61, pp. 19-30.

Nelson, C. R. and C. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series". *Journal of Monetary Economics*, No. 10, pp. 139-67.

Priestly, M. (1981): *Spectral Analysis and Time Series*. Academic Press.

Quah, D. (1992): "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds". *Econometrica*, No. 60(1), pp. 107-18.

Reinsel, G. C. and S. H. Ahn (1988): "Asymptotic Distribution of the Likelihood Ratio Test for Cointegration in the Nonstationary VAR Model". *Technical Report*, University of Wisconsin, Madison.

Runkle, D. E. (1987): "Vector Auto-regression and Reality". *Journal of Business and Economic Statistics*, No. 5, pp. 437-54.

Sargent, T. J. (1987): *Macroeconomic Theory*. Academic Press.

Schwert, G. W. (1987): "Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data". *Journal of Monetary Economics*, No. 20, pp. 1037-87.

Shapiro, M. D. and M. W. Watson (1988): "Sources of Business Cycle Fluctuations". *NBER Working Paper*, No. 2589.

Singleton, K. (1988): "Econometric Issues in the Analysis of Equilibrium Business-Cycle Models". *Journal of Monetary Economics*, No. 21, pp. 361-86.

Staiger, D., J. Stock and M. W. Watson (1996): "How Precise are Estimates of the Natural Rate of Unemployment?" Presented at the *Monetary Policy and Low Inflation Conference*, National Bureau of Economic Research.

van Norden, S. (1995): "Why is it so Hard to Measure the Current Output Gap?" *Mimeo*, Bank of Canada.

Watson, M. W. (1986): "Univariate Detrending Methods with Stochastic Trends". *Journal of Monetary Economics*, No. 18, pp. 49-75.

Watson, M. W. (1993): "Vector Autoregressions and Cointegration". *Handbook of Econometrics*, Vol. 4.

Whittaker, E. T. (1923): "On a New Method of Graduation". *Proceedings* of the Edinburgh Mathematical Society, No. 41, pp. 63-75.

Comments on: "Measurement of the output gap: a discussion of recent research done at the Bank of Canada" by Pierre St-Amant and Simon van Norden

by Florence Béranger and Gabriele Galati

This is a very interesting and well-written survey paper on different methods to estimate the output gap. The authors compare three approaches to measuring potential output and the output gap – a mechanical approach based on univariate dynamics, a pragmatic approach that augments univariate dynamics with structural relationships, and an approach that combines multivariate dynamics with long-run restrictions based on economic theory.

The discussion of the mechanical approach in the first section of the paper focuses on the popular univariate Hodrick-Prescott (HP) filter and its shortcomings. The HP filter seems ill suited as a tool for policy makers because it suffers from end-of-sample problems that make measures of the current output gap very unreliable and, moreover, it cannot be used to provide forecasts of future output gaps, which are of great interest for policy makers. Another limitation of the HP is that it associates the output gap with business cycle frequencies in the data. Instead, these may reflect the (transitional) dynamics of shocks to potential output. Also, the HP filter method produces spurious cyclicality when it is applied to integrated or near-integrated data. Finally, it performs poorly in identifying the cyclical component of time series – like those of most macroeconomic variables – that have a spectrum or pseudo-spectrum with typical Granger shape.

The paper then describes methods that augment univariate dynamics with structural relationships by adding the residuals of a structural economic relationship to the minimisation problem of the HP filter. These methods are more pragmatic but, at the same time, there are no guidelines on how to calibrate the structural relationships. They also fail to solve the end-of-sample problems and usually provide only very imprecise estimates of potential output and the output gap. The authors further describe a method (the TOFU) based on Maximum Likelihood estimates of the inverse of a simple structural relationship, but it is hard to get a clear sense of how well this method can perform in practice.

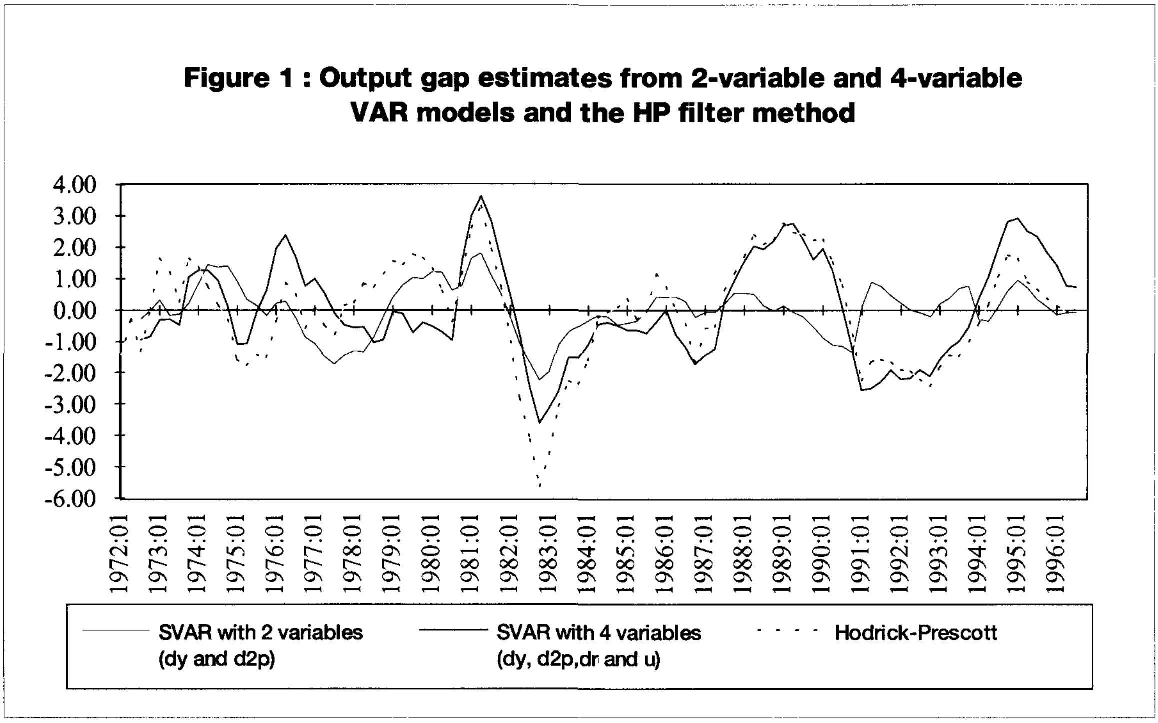
The most innovative and, in our view, interesting part of the paper describes methods of multivariate dynamics based on structural relationships. These methods use VAR models that incorporate long-run restrictions based on economic theory following the methodology of Blanchard and Quah (1989). Within this category the authors distinguish three methods: the long-run restrictions on output method (LRRO) which defines potential output as the permanent component of output and the output gap as that part of output which is due to purely transitory shocks (it should be noted that the authors are careful in avoiding the tricky issue of how to interpret permanent and transitory shocks in terms of demand and supply shocks); the Multivariate Beveridge-Nelson (MBN) method which defines potential output as the level of output that is reached after all transitory dynamics have worked themselves out; and the method proposed by Cochrane (CO) which derives identifying restrictions for a bivariate VAR model from the permanent income hypothesis. Most of the third section is devoted to a comparison of these three methods. The authors show that the LRRO outperforms the MBN and the CO methods because it lets the data determine the shape of the diffusion process of permanent shocks; i.e. it does not impose restrictions on the short-run dynamics of the permanent component of output or assume that output follows a random walk with drift.[1] In our discussion we will, therefore, focus on the LRRO method, noting some of the main advantages of this method and the problems and issues

---

[1] For the CO method this is true when consumption follows a random walk.

that it entails. It should be stressed, though, that the authors are aware of all these problems and issues, which makes their paper an excellent guide to empirical work on estimating the output gap.

A main advantage of the LRRO method is that it uses economic theory to derive restrictions on the VAR models that pin down the estimates of the output gap. An important issue concerns the choice of variables that are included in the VAR model. In order to capture the main features of real output in Canada, the authors include four variables in their VAR model – the growth rate of real output, the first difference of inflation, the unemployment rate and the first difference of the real interest rate. We can, however, think of alternative specifications, for example Blanchard and Quah's bivariate VAR model with output and unemployment, or a model that includes only output and inflation. It is important, therefore, to motivate the choice of variables. Figure 1 shows how the estimates of the output gap can be sensitive to the specification of the underlying model. It compares an estimate of the output gap for Canada derived from the four-variable model used in the paper with an estimate derived from a bivariate model with the growth rate of output and the first difference of inflation. The two measures of the output gap look quite different and have a correlation coefficient of only 0.34.[2]



Another advantage of the LRRO method is that it is flexible with respect to the properties of the variables used in the model. As the authors state, "the assumptions on the level of integration are not part of the methodology per se" and "changing these assumptions could have a significant impact on the results". The degree of differencing to be used for variables other than output (which should be I(1)) is, however, a tricky issue. The authors assume that both the real interest rate and the inflation rate have unit roots, but if these series have almost but not exactly unit roots, important information on the short-run dynamics of output may be lost by differencing them[3]. We can get a sense of whether we over-difference and lose important information by comparing the impulse

[2] The correlation coefficient rises however to 0.75 when both models are estimated using the level of inflation. Figure 1 reports also the HP filter estimate of the output gap.

[3] The estimates would still be consistent if the series are not differenced.

response functions of the bivariate and the four-variable models. Figure 2 shows that the output gap derived from the bivariate model looks very different depending on whether we difference inflation. The two series have a correlation coefficient of only 0.47. In the authors' four-variable model, however, it does not matter that much whether inflation and the interest rate are differenced or not (Figure 3).

Figure 2 : Output gap estimates in the 2-variable model with inflation and the first difference of inflation

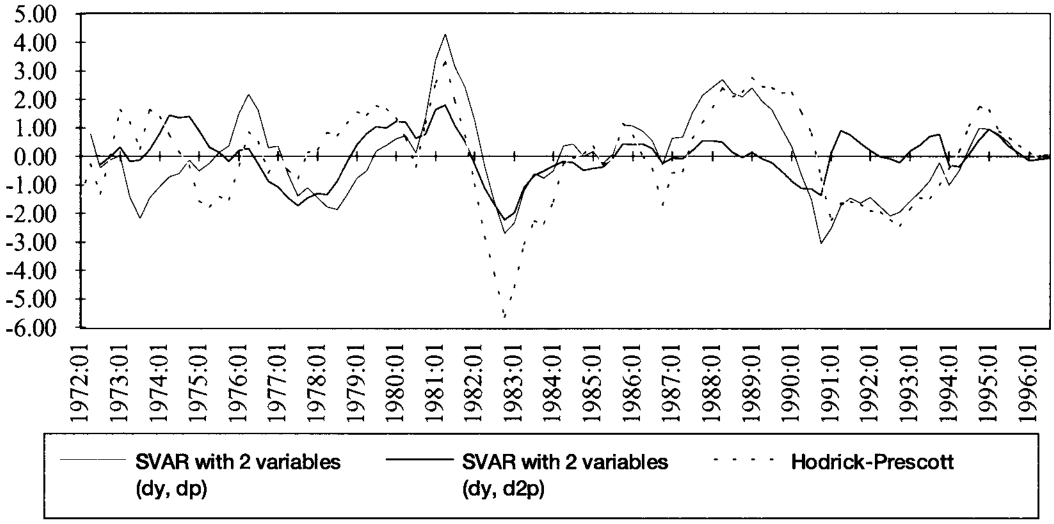
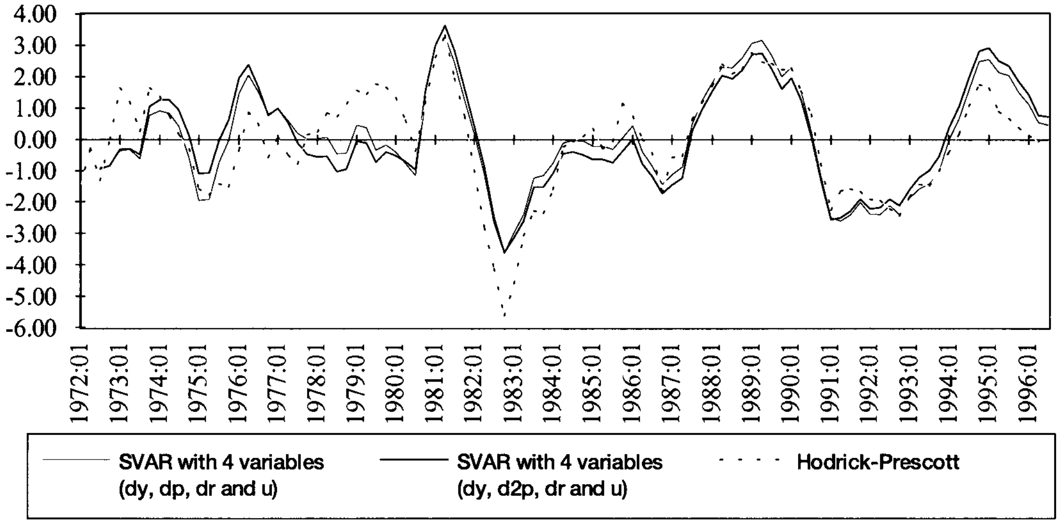


Figure 3 : Output gap estimates in the 4-variable model with inflation and the first difference of inflation



Finally, an advantage of the LRRO method over the mechanical approaches is that it provides estimates of the current as well as of the expected future output gap. However, these estimates are generally very uncertain. There is a trade-off in choosing the number of lags in a VAR model between keeping it low to reduce the uncertainty of the estimates and keeping it high to avoid the bias in the estimation of the structural components. Still, Figure 15 in the paper shows several episodes where the LRRO output gap estimate for Canada is significantly different from zero in several instances. The authors also find that their results are robust to using 6 or 10 lags instead of 8.

References

Blanchard, Olivier J. and Danny Quah (1989): "The Dynamic Effects of Aggregate Demand and Supply Disturbances". *American Economic Review*, No. 79, pp. 655-673.

An empirical assessment of the link between the output gap and inflation in the French economy[1]

John Baude and Gilbert Cette

Introduction

There may be many reasons for making a diagnosis of the position of an economy in the economic cycle. Such a diagnosis makes it possible, for example, to assess the economy's growth potential, potential inflationary pressures and the structural position of government finances. This paper addresses the second of these three concerns. For the French economy it attempts to give an empirical analysis of the relationship between the position in the cycle, as determined by various indicators of the output gap, and inflation. It draws on previous work on the same subject, such as by Giorno et al. (1995) on estimations of the output gap and by Turner (1995) on the link between these indicators and inflation.

The choice of the methods used to estimate output gaps and their link with inflation was based on three criteria which we felt to be essential and which played a major role in our final decision.

1. The methods had to be *reproducible* by any economist using the same statistical data. This meant that we were unable to include so-called "expert" opinion.
2. They had to be *easy to apply* to different industrialised countries using standardised databases. For that reason it was important to limit the volume of data used in the calculations.
3. They had to *produce quick results*, so as to be systematically repeatable at low costs whenever the database is updated or changed.

These three criteria lie behind the original and specific features of the present study, as regards both construction of the output gap indicators (e.g. in the measurement of the capital stock and the estimation of an equilibrium rate of unemployment) and determination of the link between inflation and these indicators. The methods are deliberately crude. They cannot claim to provide a detailed diagnosis of the position of economies in the cycle or of the link between this position and inflation. The results must necessarily be compared with results obtained using other methods. It should also be borne in mind that this study is a reflection of work in progress. It does not in any way represent an official position of the Banque de France.

Section 1 looks at the methods used for estimating the output gap and their results, while the links with inflation are analysed in Section 2.

1. Estimations of the output gap[2]

Economists have put forward a variety of alternative methods for diagnosing an economy's position in the economic cycle. They may be based on a single variable, in which case only

[1] This paper reflects work in progress at SEMEF. It does not in any way represent an official position of the Banque de France. Some of the data processing for this study was carried out by Laurent Baudry, Lydie Gomez and Béatrice Saes-Escorbiac.

[2] This is a summary of more detailed arguments contained in Cette (1997).

data concerning output (GDP if the study is carried out at an aggregate level) are used. Alternatively they may be based on multiple variables, in which case the approach may be structural (e.g. based on explicit production functions) or non-structural (e.g. based on VAR models).

The *time horizon may be more or less distant.* Within this context, the methods may also assume a greater or lesser degree of flexibility with regard to factors of production. In single-variable studies, the choice of time horizon consists of "restricting", to a greater or lesser extent, the acceptable variations of potential output. In multiple-variable studies, it consists of symmetrically "restricting", to a greater or lesser extent, the acceptable modifications of the variables explaining potential output. Let us take structural methods as an example. If we consider the very short term, stocks of factors of production may be regarded as fixed and the gap between actual and potential output is explained simply by degrees of factor utilisation. If we consider the longer term, factor stocks may be regarded as adjustable, within more or less explicit limits. For example, for the labour factor this could be the immediately available manpower (i.e. the labour force) or potentially mobilisable manpower (i.e. the potential labour force after taking cyclical declines in participation rates into account). Over the very long term, stocks of the two factors may even be regarded as entirely adjustable (e.g. by recourse to immigration for the labour factor), in which case potential growth becomes indeterminate.

This variety of methodological options explains why so many estimations and methods have been proposed in the academic literature since the pioneering work of Okun (1962), who proposed a simple linear relationship between the deviation of unemployment from its natural level and the deviation of output from its potential level. A wide-ranging critical review of the literature is contained in Cour, Le Bihan and Sterdyniak (1997), referred to subsequently as CLS (1997). Likewise, there is no evidence that similar players in different countries (e.g. central banks) favour one particular methodological approach over another, or that different players (e.g. international organisations) favour one particular approach for the same purpose (e.g. formulating medium-term growth scenarios).

There is no shortage of examples of this diversity of methods. The US Federal Reserve uses a method based on Okun's law (cf. Kahn (1996)). The Bank of England uses a number of single- and multiple-variable methods (cf. Fisher, Mahadeva and Whitley (1997)). The Bundesbank uses a structural method in which the output gap is explained by degrees of factor utilisation alone (cf. Bundesbank (1995) and Westermann (1997)). The same pattern is to be found among international organisations, which use such estimations to formulate medium-term growth scenarios. The European Commission uses a single-variable method based on a smoothing of GDP (cf. Ongena and Röger (1997)). The IMF uses a structural method in which the output gap is explained by degrees of factor utilisation and by the gap between the unemployment rate and an estimation of the NAIRU (cf. De Masi (1997)). The OECD uses a number of single- and multiple-variable methods. Among the latter, particular emphasis is placed on a structural approach in which the output gap is explained by degrees of factor utilisation and by the gap between the unemployment rate and an estimation of the NAWRU (cf. Giorno, Richardson, Roseveare, Van Den Noord (1995), subsequently referred to as GRRV (1995), and Giorno and Suyker (1997)).

In the context of this wide diversity of methods, the approaches described here, which continue earlier work (cf. Villetelle (1994)), have a distinctly pragmatic cast. Rather than preferring one estimation to another, we have defined three alternative calculations of potential output and, as a corollary, of the output gap. Our diagnosis is enriched by a comparison of the different results. Two of the methods are single-variable methods. The third is a multiple-variable method with a structural component in which the output gap is explained by the degree of factor utilisation and by the gap between the actual and equilibrium rate of unemployment estimated in a specific way. This third method is the only one to display any original features, as the equilibrium rate of unemployment is defined as the rate which would stabilise firms' profit ratios in the short term. Several calculations of the profit ratio are envisaged. One of them takes into account the flow of firms' net interest payments. This makes it possible to include financial considerations when determining the equilibrium rate of unemployment and hence potential output and the output gap.

Notation and data sources

PIB	Gross domestic product (GDP) or output
PIBL	Smoothed output
PIBT	Trend-based output
PIBPi	Potential output. Four estimations are proposed corresponding to the profit ratios *TM*1 to *TM*4
EPIB	Output gap. The suffix *L*, *T* or *Pi* indicates the output indicator used to define the gaps. Hence *PIB* = *PIBL* + *EPIBL* = *PIBT* + *EPIBT* = *PIBPi* + *EPIBPi*
N	Domestic employment
K	Real productive fixed capital employed in productive activity
E or *NE*	as a suffix of the variables *PIB*, *N* or *K*, they indicate that these values relate respectively to non-financial firms only or to the rest of the economy. Hence *PIB* = *PIBE* + *PIBNE*; *N* = *NE* + *NNE* and *K* = *KE* + *KNE*
$f(t)$	Hicks-neutral technological progress in non-financial firms. Hence: $PIBE = \alpha(KE_{-1}) + (1-\alpha)NE + f(t)$, where $0 \leq \alpha \leq 1$
TMi	Firms' profit ratio. Four profit ratios are calculated: a current profit ratio *TM*1 where gross operating surplus at factor cost (balance *N*2 of the operating account for firms in the national accounts) is related to the value added at market prices (balance *N*1 of the production account for firms), a profit ratio *TM*2 at market prices, a profit ratio *TM*3 at factor cost and a profit ratio *TM*4 at factor cost excluding net interest charges
α_j	Weighting coefficient for capital in the Cobb-Douglas function representing firms' combination of factors of production. The value of the coefficient is the average of the actual profit ratio between 1970Q1 and 1996Q2. Four values have been calculated: $\alpha1 = 0.256$ for *TM*1, $\alpha2 = 0.312$ for *TM*2 and $\alpha = 0.275$ for *TM*3. As *TM*4 is not really a profit ratio, it is assumed that $\alpha4 = \alpha3$
TC	Unemployment rate
TCi^*	Equilibrium rate of unemployment. Defined by the equation: $TCi^* = TC - \beta^{-1}\Delta TMi$, where $\beta = 0.5$. Four equilibrium unemployment rates are calculated, corresponding to the four profit ratios defined above.
π_N	Productivity per capita: $\pi_N = PIBE\,/\,NE$
δ	Retirement rate of capital goods
TUA	Production capacity utilisation rate (including recruitment)
EBE	Gross operating surplus : $EBE = P.PIBE - W.NE$
Pc	Consumer prices
P	Price of value added
Pm	Price of imports of goods and services
Pe	Average consumer price index of France's nine leading trading partners adjusted for exchange rates. The weighting given to each is the share of imports of goods and services from that country
W	Wages per capita
TI	Nominal interest rate
*	As a suffix of a variable, indicates its equilibrium level
Δ	In front of a variable, indicates its variation from one period to another
.	Above a variable, indicates its growth rate
l	As a suffix of a variable, indicates its smoothing
–	Above a variable, indicates its average
L	Lagging operator
$\phi(L)$	Lagging operator polynomial
*AV*1	Other variables affecting the rate of wage growth
*AV*2	$\Phi(L)\dot{P}_c - \dot{P}$

Lower-case variables correspond to their logarithm

Data sources: except where otherwise stated, all data used in this study are drawn from the quarterly national accounts

The chosen methods are deliberately simple. They cannot claim to provide a detailed diagnosis of the position of economies in the cycle. The results should rather be compared with those derived from other methods. We shall begin by describing the three methods used (Section 1.1) before discussing the main results (Section 1.2).

1.1 The methods used

The first estimation of the output gap (written as *EPIBL*) is based on a smoothing of the output logarithm using the Hodrick-Prescott filter. The usual standard value for the smoothing parameter (λ = 1,600) is applied to the quarterly data used here. We shall not go into the very numerous limitations of this method since they are discussed in detail in the literature (cf. for example Allard (1994), CLS (1997) and Berger and Teil (1996)). The main advantages of the method are that it is quick, easy to use and reproducible and that the results are easy to interpret.

The second method (written as *EPIBT*) is based on estimating a determinist output trend in order to obtain trend-based output, taking into account possible breaks in the trend when such are suggested by analysis of the residuals. The method is standard (cf. for example INSEE (1995)) and it calls for determination of the trend and possible breaks. Our estimations have been made using an algorithm designed by Berger and Teil (1996) which determines endogenously the most significant combination of significant breaks in the trend. Although this method is relatively sophisticated from a technical standpoint, it remains open to the usual criticisms made of any determinist approach (cf. for example Berger and Teil (1996)). It has the same advantages as the smoothing method and the fact that it is entirely reproducible is a particularly important feature for estimations of this type.The estimations carried out for France over the period 1960-95 show two breaks in the output trend, in 1973Q3 and 1980Q2. Accordingly, the underlying growth rate of French GDP is approx. 2.0% since 1980.

A third method (written as *EPIBP*) is based on a structural approach to the calculation of potential output, which combines the choice of factors of production with determination of an equilibrium rate of unemployment (*TCE*). The estimation is carried out in the following stages.

i) First, it is assumed that only non-financial firms are endogenous, whereas the rest of the economy is exogenous.

ii) Our estimation of firms' fixed capital stock is not derived from national accounts, which are based on assumptions that are inevitably fragile and that differ widely from one country to another. Indeed, we have estimated the capital stock, assuming the sudden death of capital goods and an average lifetime of 12 years. The sudden death hypothesis has only a marginal effect on the profile of the statistical series derived in this way (cf. Maddison (1993)). The assumption that capital goods have a lifetime of 12 years (48 quarters) is based on estimations carried out on large samples of French firms (cf. Cette and Szpiro (1988)).[3]

iii) Our specification of firms' choice of factors of production is based on a Cobb-Douglas function with constant returns to scale. It further assumes that technological progress $f(t)$ is Hicks-neutral and that factors of production are limited to the stocks of labour *NE* and capital *KE*; the mobilised capital stock *KE* at quarter t is the fixed capital stock at the end of the previous quarter. This gives:

$$pibe = \alpha(ke_{-1}) + (1-\alpha)ne + f(t), \text{ where } 0 \leq \alpha \leq 1 \quad (1)$$

As usual, estimation of equation (1), with the additional assumption of a deterministic trend (with possible breaks) for the effects of technological progress $f(t)$, gives aberrant results for the α parameter (cf. Berger and Teil (1996)). Thus, the value of α was made equal to the average of firms' actual profit ratio *TM* over the period 1970Q1 to 1995Q4. Four profit ratios were calculated: a current profit ratio *TM*1 where gross operating surplus at factor cost (balance *N*2 of the operating account for firms in the national accounts) is related to value added at market prices (balance *N*1 of the production account for firms), a profit ratio *TM*2 at market prices, a profit ratio *TM*3 at factor cost, and a profit

[3] For a more detailed consideration of these matters, see Cette (1994).

ratio *TM*4 at factor cost excluding net interest charges.[4] A comparison of *TM*3 and *TM*4 shows the impact of changes in interest charges on the equilibrium rate of unemployment and the output gap. As *TM*4 does not really correspond to the share of capital in the primary distribution of income (it is not the complement of the share of labour costs in value added), only the first three profit ratios were used for the α coefficient. The α parameters calculated from these profit ratios are as follows: $\alpha 1 = 0.256$ for *TM*1, $\alpha 2 = 0.312$ for *TM*2 and $\alpha 3 = \alpha 4 = 0.275$ for *TM*3.

Equation (2) below was then used to calculate the Solow residual for technological progress $f(t)$ from each of the three profit ratios, here given the index j:

$$f(t)_j = pibe - \alpha_j(ke_{-1}) + (1 - \alpha_j)ne \qquad (2)$$

iv) The fourth stage consisted of calculating the equilibrium rate of unemployment *TC**, which may be defined as the rate which implies no acceleration of wages (NAWRU) or prices (NAIRU) or no change in the profit ratio. We used the latter definition because it is more effective than the other two in limiting the difficulties of satisfactorily including the effects of changes in the terms of trade. This method for calculating the equilibrium rate of unemployment has the added advantage of being quick and easy to use (also in the context of macroeconomic forecasts) since the only data required is a series of non-financial firms' profit ratios. The equilibrium rate of unemployment is thus distinguished by the absence of short-term inflationary (or disinflationary) wage pressures due to a conflict between wages and profits in the distribution of primary income.

To illustrate the method, the equilibrium rate of unemployment *TC** is defined assuming that the change of the smoothed profit ratio *TM*1 (using a Hodrick-Prescott filter with $\lambda = 100$)[5] between two dates is proportional to the difference between the smoothed unemployment rate *TC*1 (using a similar Hodrick-Prescott filter) and the equilibrium rate *TC**:

$$\Delta TM1 = \beta\left(TC1 - TC^*\right) \qquad (3)$$

Hence: $TC^* = TC1 - \frac{1}{\beta}\Delta TM1$, or: $TC^* - TC = (TC1 - TC) - \frac{1}{\beta}\Delta TM1$, where $\beta > 0$.

This link between changes in the profit ratio and the gap between the actual and equilibrium rate of unemployment is based on the method used in OECD studies to calculate the NAWRU (cf. GRRV (1995)), which links wage acceleration to the gap between the actual and equilibrium rate of unemployment. The method was originally put forward by Elmeskov (1993) and Elmeskov and Macfarlan (1993). We show in Annex 1 that under certain assumptions equation (3) can be deduced from a simplified price-wage loop. This calculation of the equilibrium rate of unemployment differs from previous, more sophisticated estimations of the NAIRU carried out at the Banque de France, based on structural or reduced price-wage loops (cf. Jackman and Leroy (1995)). The equilibrium rate of unemployment estimated here corresponds to a short-term approach. Although the calculation is based on variables that have been previously smoothed, smoothing is carried out over short periods and thus does not in any way correspond to a calculation of the structural rate of unemployment (cf. Layard, Nickell and Jackman (1991)).

---

4 The calculation of non-financial firms' profit ratios was adjusted for the impact of the growth of the wage-earning class. For a more detailed consideration of these matters, see Cette and Mahfouz (1996).

5 The choice of short-period smoothings ($\lambda = 100$) is due to the fact that longer-period smoothings (e.g., $\lambda = 1{,}600$) cause the effects of the first oil shock to be reflected in the equilibrium unemployment rate even before 1973, which seems absurd to say the least.

The gap between the actual and equilibrium rate of unemployment combines two effects: the smoothing of the unemployment rate and changes in the smoothed profit ratio (equation (3)). The β parameter affects only the amplitude of the second effect and not its sign. The effect is positive before the first oil shock (the profit ratio improves), then negative until the early 1980s (the profit ratio deteriorates), then positive again until the early 1990s and has remained positive in recent years. There is little change in the behaviour of the parameter according to the different profit ratios used. In the recent period, the fall in firms' net interest charges (due to lower average interest rates and firms shedding debt) has caused the equilibrium rate of unemployment calculated from the profit ratio (*TM*4) to diverge significantly from the three others.

As explained in Annex 1, we decided to set $\beta = 0.5$ as a standard value. We then used equation (3) to calculate four equilibrium unemployment rates (*TC*1* to *TC*4*) from the four previously defined profit ratios (*TM*1 to *TM*4).

v) We then calculated non-financial firms' potential employment (*NE**) as the difference between total potential employment in France and employment excluding non-financial firms, smoothed using a Hodrick-Prescott filter (with $\lambda = 1{,}600$). Total potential employment in France was calculated by applying the previously calculated equilibrium rate of unemployment to the labour force (according to ILO definitions) smoothed using a Hodrick-Prescott filter (with $\lambda = 1{,}600$). Four series of non-financial firms' potential employment were calculated in this way, corresponding to the four series of equilibrium rates of unemployment.

vi) The next stage was to calculate non-financial firms' potential value added by applying the production function represented by equation (1) to potential quantities of factors. The potential effect of technological progress ($f(t)^*$) was estimated by its smoothed value using a Hodrick-Prescott filter (with $\lambda = 1{,}600$), the potential capital stock by its actual level ($K^* = K$), and non-financial firms' potential employment as described above. Four series of non-financial firms' potential value added were calculated in this way, corresponding to the four series of non-financial firms' potential employment and of the corresponding values for the factor weighting parameter.

vii) Finally the level of potential output (*PIBP*) was calculated by adding non-financial firms' potential value added to the value added of the rest of the economy smoothed using a Hodrick-Prescott filter (with $\lambda = 1{,}600$). Four series of potential output (*PIBP*1 to *PIBP*4) and of potential output gaps (*EPIBP*1 to *EPIBP*4) were calculated in this way, corresponding to the four series of non-financial firms' potential value added.

Clearly, this third approach, like the other two, has a number of weaknesses relating to the various simplifying assumptions included in the calculation. Like all estimations of potential output, the ones proposed here are crude and, inevitably, relatively imprecise. Thus, the indicators should be considered more for what they show, when they are consistent with each other or when their divergence can be interpreted, rather than for their actual levels.

1.2 Main results

We should emphasise that these estimations have been carried out using data from the national accounts, supplemented by Banque de France forecasts[6] for the French economy. Consequently the results for recent years may differ significantly from those of other studies based on similar methods but using other forecasts to extend historical data series.

i) The estimated potential growth rates lead to a fairly common diagnosis: potential output growth in France slowed considerably over the period, falling from approx. 4.0-5.0% (depending on the indicators) before the first oil shock to approx. 1.5-2.0% in recent years. The two

6 These forecasts estimate average annual output growth in France at 1.1% in 1996, 2.3% in 1997 and 2.8% in 1998.

main slowdowns occurred after the first oil shock and in the early 1990s (see Table 1). As this broad analysis has already been developed in numerous other papers on the subject (cf. for example GRRV (1995), Giorno and Suyker (1997) and Bouthevillain (1996)) we shall not go into further detail here.

Table 1

Potential and actual growth of French output (in percentages)

Output indicator	**Abbrev.**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**
Actual	*PIB*	4.5	4.3	2.5	0.8	1.2	-1.3	2.8	2.2	1.1*
Smoothed	*PIBL*	3.0	2.8	2.3	1.8	1.3	1.2	1.3	1.5	1.7
Trend-based	*PIBT*	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Current potential	*PIBP*1	3.2	3.0	2.4	1.7	1.0	0.9	1.3	1.7	1.8
Potential at market prices	*PIBP*2	3.2	3.0	2.4	1.7	1.1	1.0	1.4	1.7	1.8
Potential at factor cost	*PIBP*3	3.2	3.0	2.4	1.7	1.0	1.0	1.4	1.7	1.8
Potential at factor cost excl. interest charges	*PIBP*4	3.0	2.8	2.4	1.6	1.1	1.2	1.5	1.6	1.7

* Forecast.

For two main reasons the low potential growth rates of recent years should not be regarded as definitive. First, recent figures for real growth have not yet been finalised and are thus liable to change. Secondly, estimations of potential output for recent years could be modified (on account of the smoothing or trend adjustment methods) according to actual growth in 1996 and subsequent years. Thus, the same estimation methods will give higher potential growth rates for the years 1993-95 if actual growth is higher in subsequent years.

Table 2

Gaps between potential and actual output in France (in percentages)

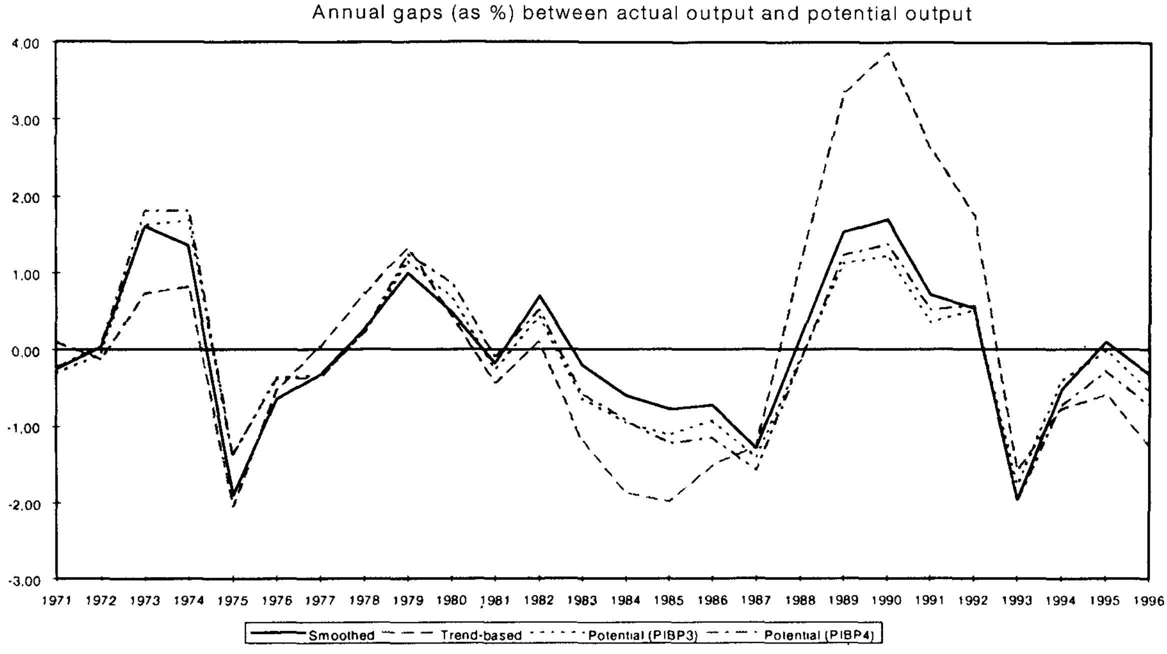
Potential output indicator	**Abbrev.**	**1988**	**1989**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**
Smoothed	*EPIBL*	0.1	1.5	1.7	0.7	0.5	-2.0	-0.5	0.2	-0.5
Trend-based	*EPIBT*	0.4	3.4	3.9	2.7	1.8	-1.5	-0.7	-0.5	-1.5
Current potential	*EPIBP*1	-0.1	1.1	1.2	0.3	0.5	-1.7	-0.3	-0.2	-0.5
Potential at market prices	*EPIBP*2	-0.1	1.1	1.2	0.3	0.5	-1.9	-0.5	-0.1	-0.6
Potential at factor cost	*EPIBP*3	-0.1	1.1	1.2	0.3	0.5	-1.8	-0.4	-0.1	-0.6
Potential at factor cost excl. interest charges	*EPIBP*4	-0.2	1.2	1.4	0.5	0.5	-2.0	-0.7	-0.2	-0.8

ii) There is nothing unusual about the variations in output gaps and the position in the cycle to which they correspond (cf. for example the studies cited above), so we will not comment on them in any further detail here (see Table 2 and Chart 1). The fact that output gaps are relatively small in the last two years can be attributed to the same reasons as those given earlier concerning potential output growth.

A consideration of the links between the various output gaps, the gaps between actual and equilibrium rates of unemployment and the capacity utilisation rate,[7] gives the following results (cf. Annex 3, Table A3-1).

[7] In this case manufacturing firms' production capacity utilisation rate including recruitment (*TUA*) measured by INSEE from its quarterly economic survey.

Chart 1



The smoothed output gap (*EPIBL*) and the four output gaps (*EPIBP*1 to *EPIBP*4) from a structural approach entirely correlate with each other (the correlation is of the order of 99%). The correlation with the trend-based output gap (*EPIBT*) is smaller (75-80%). The output gaps also correlate with the four gaps between actual and equilibrium rates of unemployment, and the correlation is strong (60-70%) for the four "structural" gaps (*EPIBP*1 to *EPIBP*4) and less strong (40-50%) for the other two (*EPIBL* and *EPIBT*). All output gaps correlate fairly strongly (75-85%) with the production capacity utilisation rate, though the correlation is less strong (60%) for the trend-based output gap.

The four gaps between actual and equilibrium rates of unemployment (*TC*1*–*TC* to *TC*4*–*TC*) correlate entirely with each other (the correlation is of the order of 99%), while the correlation with the production capacity utilisation rate is much less strong (20-30%). This suggests that output gap indicators do indeed synthesise the pressures on both goods and labour markets. It also suggests, however, that the indicators of labour market pressures provide different information from the indicators of pressures in the goods market.

The econometric results that explain output gaps in terms of the pressures in the two markets show that a gap of one point more (less) between actual and equilibrium rates of unemployment increases (reduces) the output gap by 0.8-0.9 of a point, and that one point more (less) on the production capacity utilisation rate increases (reduces) the output gap by 0.3 of a point. The effect of each pressure variable seems fairly robust with regard to both the presence of the other and to the method used to estimate the output gap. These results confirm that the potential output gap indicators developed here reflect pressures exerted in the goods and labour markets simultaneously.

Two results relating to the level of estimated output gaps deserve particular attention.

First, the trend-based output gaps (*EPIBT*) are larger over the last two years than the other output gaps. This result is simply due to the fact that the average output trend since 1987 (when the second and last break in the trend occurred) is higher than the growth rates of the other potential output indicators which were dampened by low actual rates of GDP growth in the most recent years.

Secondly, the potential output gap at factor cost excluding net interest charges (*EPIBP4*) was 0.2 of a point larger in 1996 than the potential output gap at factor cost (*EPIBP3*). This difference is due to the fall in firms' net interest payments (expressed in points of value added) in recent years as they have shed debt and as interest rates have fallen (cf. Cette and Mahfouz (1996)). This fall in interest charges has influenced the comparative trends in the two output gaps through its impact on estimations of the equilibrium rate of unemployment (see below).

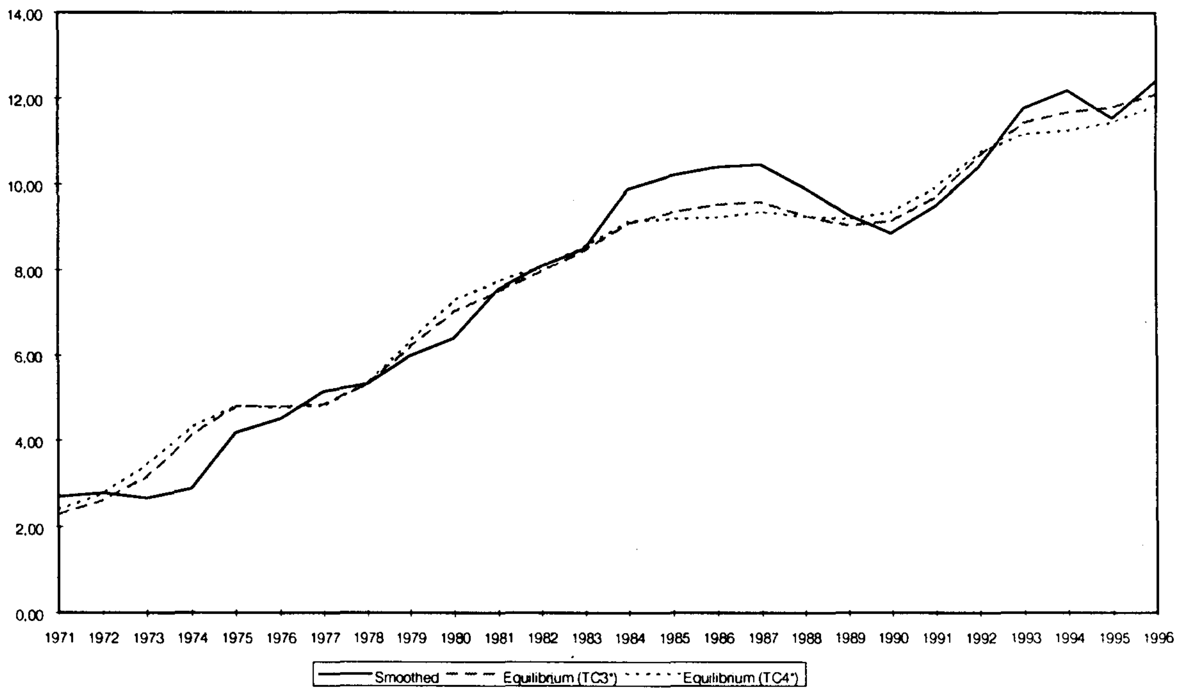
Table 3

Gaps between the actual and equilibrium rate of unemployment in France (in percentage points)

Equilibrium rate of unemployment indicator	Abbrev.	1988	1989	1990	1991	1992	1993	1994	1995	1996
Current	*TC1**	0.6	0.2	-0.3	-0.2	-0.2	0.3	0.4	-0.4	0.1
At market prices	*TC2**	0.7	0.3	-0.3	-0.2	-0.2	0.4	0.6	-0.2	0.3
At factor cost	*TC3**	0.7	0.3	-0.3	-0.2	-0.2	0.4	0.5	-0.3	0.2
At factor cost excl. interest charges	*TC4**	0.7	0.1	-0.5	-0.4	-0.3	0.6	0.9	0.1	0.5

Chart 2

Unemployment rate (annual average, as %)



iii) Variations in the short-term equilibrium rate of unemployment diverge relatively little from variations in the actual unemployment rate (Table 3 and Chart 2), because the former do not correspond to unemployment rates below which inflationary pressures would inevitably appear, but merely to unemployment rates which (under various simplifying assumptions) would allow various profit ratios to stabilise in the short term, taking actual wage pressures on the labour market into account. In other words, these equilibrium unemployment rates are liable to fall substantially in future periods if the situation on the labour market can be improved without a parallel fall in firms' profit ratios, caused either by wage pressures or by increased competitive pressures on price formation.

The gap between actual and equilibrium unemployment rates is distinctly negative from the first oil shock until the end of the 1970s (firms' profit ratios fell significantly), but positive throughout the 1980s (firms' profit ratios rose significantly). It was between 0.5 and 1 of a point in 1994 and between 0.1 and 0.5 of a point in 1996. Because of the fall in firms' net interest payments in recent years, the equilibrium rate of unemployment which takes this element into account (*TC*4*) is lower than those that do not (e.g. *TC*3*).

2. A link between the output gap and inflation

2.1 The expected link between the output gap and inflation

What influence might an economy's position in the economic cycle have on inflation? The Phillips-curve equation alone cannot provide the answer, because the continuing upward drift of unemployment and the absence from the equation of any indicator of pressures on the goods market mean that it cannot reflect the short-term inflationary pressures in the economy, the time horizon on which we have focused. We, therefore, substituted a cyclical indicator of inflationary pressures in the goods market. In this way we, alternatively, tested the output and the capacity utilisation gaps. The former, whether derived from a smoothed trend (*EPIBL*) or by calculating potential output (*EPIBP*3), is supposed to synthesise pressures on all goods and labour markets, and the latter, pressures in the goods market alone.

The output or capacity utilisation gap may have a dual influence on inflation. At times of economic recovery, for example, there is a certain time lag before the effect is reflected in staffing levels and the capital stock. Firms, therefore, use factors of production more intensively before recruiting and investing. The resulting decline in unit costs attenuates price growth at an unchanged marginal profit rate. At the same time, however, growing pressures in the goods market and increasing wage claims as conditions in the labour market improve cause prices to accelerate, still assuming that margins remain stable. The combined effect, summarised in a single, reduced equation, is thus theoretically indeterminate, though most studies conclude that the overall effect is positive. Inflation may accelerate even though output has not reached its potential level. The faster output gap is absorbed, the stronger this factor affects inflation, which is then liable to accelerate (Turner (1996)).

Price acceleration during the upper phase of the cycle may be more pronounced than the deceleration observed when the output gap is negative (cf. Turner (1995), Clark et al. (1996)). This possible assymetry, not tested in this study, is based on the Keynesian idea of an inflected supply curve that is almost vertical beyond the level of potential output.

2.2 The estimation period and the equation

A cyclical indicator of activity reflects cyclical movements in inflation but cannot explain a structural change such as the marked price deceleration over the periode 1982-86. Consequently it is preferable for the estimation periode not to include this transition phase. Being a time of instability, it could blur any measurement of the effect that the output or capacity utilisation gap might have on inflation. The estimation period using quarterly data, which is, therefore, discontinuous, included two phases corresponding to periods of high then low inflation. The first period begins in 1973 and stops in the first quarter of 1982, before the wage and price freeze was decided and put into effect. The second, which follows the last devaluation of the French franc, begins in the second quarter of 1987 and continues until the end of 1994.[8] Choosing a truncated period like this means that, for each of the two phases, price formation is assumed to have remained stable, give or take a constant. The assumption underlying this choice is implicit in the fact that there is only one break in the series, which affects the constant and not the behaviour of the explanatory variables. Statistical tests clearly show a change in average inflation (Ponty (1997)) which, although difficult to situate precisely, seems to have taken place during the price deceleration phase. It may also correspond to a far-reaching change in expectations following a change in economic policy. Recent research by Fisher et al. (1997), which is not based on a truncated period as is the case here, takes official price growth objectives into consideration in expectations of inflation.

Inflation in France does not result solely from internal pressures related to activity, whether measured by the output or the capacity utilisation gap. It also depends on variations in prices outside France. Growth of the import deflator can take this into account, but it has the disadvantage of including the margin behaviour of foreign exporters on the French market, largely determined by inflation in France. Consequently we shall see whether another indicator, namely the average inflation adjusted for the exchange rate of France's main trading partners, might no be more relevant. Lastly, one or more autoregressive terms will be included in order to correct any autocorrelation of residuals. The equation, in its most general form, is as follows:

$$\dot{p}_c = \sum_i a_i \dot{p}_{c-i} + \sum_j b_j GAP_{-j} + \sum_k c_k \dot{p}_{m-k} + d + d' 1_{[87Q2-94Q4]}$$

GAP : indicator of internal pressures:

output gap ($EPIBL$ or $EPIBP3$)

capacity gap $\left(ECAP = TUA - \overline{TUA}\right)$

$1_{[87Q2-94Q4]}$: dummy corresponding to the period 1987Q2 to 1994Q4

2.3 Results

Over the survey period, all the internal pressure indicators exert a positive and significant influence after two quarters. Only the capacity utilisation gap required smoothing[9] (see Table 4). The influence proves to be relatively unaffected by the choice of the variable designed to reflect external inflationary pressures. However, we preferred to include changes in the import deflator in subsequent regressions because the regressions seemed to be of higher quality from a statistical standpoint (see Tables A3-1 to A3-4 in the Annex).

The medium-term elasticity of the output gap (0.14) is twice as large as that of the capacity utilisation gap (0.07). However, the two indicators displayed such highly dissimilar variations of amplitude that their elasticities are not comparable as they stand. But once the indicators have been centred on the mean and reduced by the standard deviation, they prove to be identical and close to 0.15 (see Table 5 and, in the Annex, Table A3-7). Thus, a one-point centred and reduced gap maintained for one year would entail price acceleration of 0.6 point. The absence of any speed limit effect in the regressions is due perhaps to the fact that the output gaps used in this study are short-term gaps which are quickly absorbed. The medium-term impact of an acceleration of imported inflation on price growth in France seems to be relatively weak. When import price growth accelerates by one point quarter-on-quarter, inflation in France increases by only about 0.1 point.

---

8 Quarters subsequent to the fourth quarter of 1994 were not included in the estimation because "extremity effects" can adversely affect the reliability of the output gap.

9 Four quarters were smoothed as follows : $TUA_{\text{smoothed}} = (TUA + 1.5TUA_{-1} + 1.5TUA_{-2} + TUA_{-3})/5$.

Table 4

Measurement of inflation with alternative indicators of internal pressures (*IPI*) and with growth of the import deflator

IPI	$\dot{P}_{-1}$	IPI_{-2}	$\dot{P}_m$	**constant (x100)**	**break in constant (x100)**	R^2	σ **(x100)**
Smoothed ouput gap	0.29 (3.2)	0.09 (2.0)	0.09 (5.9)	1.57 (6.2)	-1.14 (-6.9)	0.88	0.40
Potential output gap	0.28 (3.0)	0.10 (2.2)	0.08 (5.9)	1.59 (6.3)	-1.14 (-5.5)	0.88	0.40
Capacity utilisation gap (smoothed)	0.31 (3.4)	0.05 (2.1)	0.09 (6.0)	1.56 (6.3)	-1.15 (-5.5)	0.88	0.40

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Table 5

Medium-term elasticities between inflation and the ouput (or capacity utilisation) gap and imported inflation

Medium-term effect	**Impact of one point on the output gap**	**Impact of one point on the capacity gap**	**Impact of one point on the *centred and reduced* output or capacity gap**	**Impact of one point on imported inflation**
Impact over 1 quarter	0.14	0.07	0.15	0.12
Impact maintained over 1 year	0.6	0.3	0.6	0.5

These effects appear to be stable from one phase to another: most breaks in behaviour prove to be insignificant. Thus, the output or production capacity gap and imported inflation appear to have the same influence on price growth in periods of both high and low inflation (see Tables A3-5 and A3-6 in the Annex). However, some results are sensitive to the definition of the estimation period. The effect of the smoothed or potential output gap no longer appears to be significant when the years 1973-75 relating to the first oil shock are left out. But the robustness of the equations would need to be further assessed by making a sweep of this kind over the entire period. Dynamic simulation[10] of the equations over each of the two phases sheds some initial light (see Charts 3 to 5). The main swings in price growth seem to be fully taken into account until the early 1990s. However, actual and simulated curves diverge somewhat thereafter: during the period 1993-94, when the capacity utilisation gap was used as an indicator of inflationary pressures; and more clearly during the period 1995-96 when the output gap was used.[11]

[10] Simulated price growth, shown in graphic form, is calculated quarter-on-quarter. The curve is smoother and easier to interpret than the quarterly growth curve simulated directly from the equation.

[11] Simulated price growth for the years 1995-96 (i.e. beyond the estimation period) could not take account of every possible change in the manner of price formation occuring during those years. Moreover, there may be "extremity effects".

Chart 3

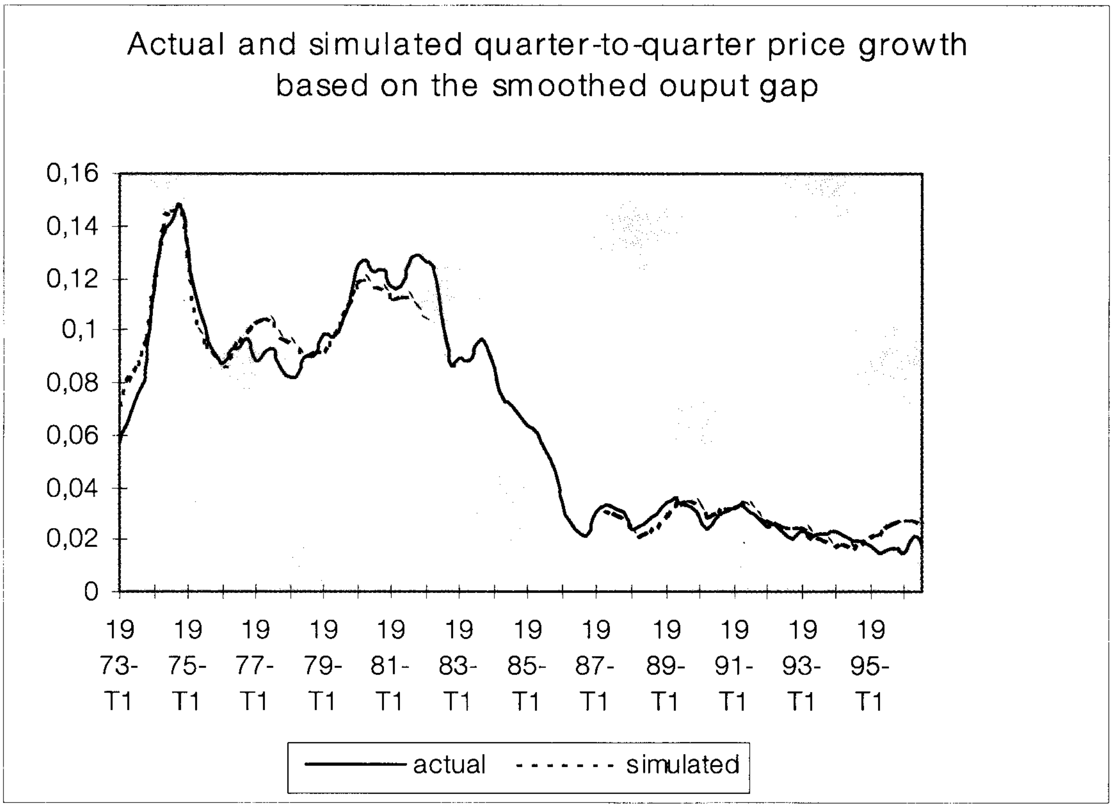


Chart 4

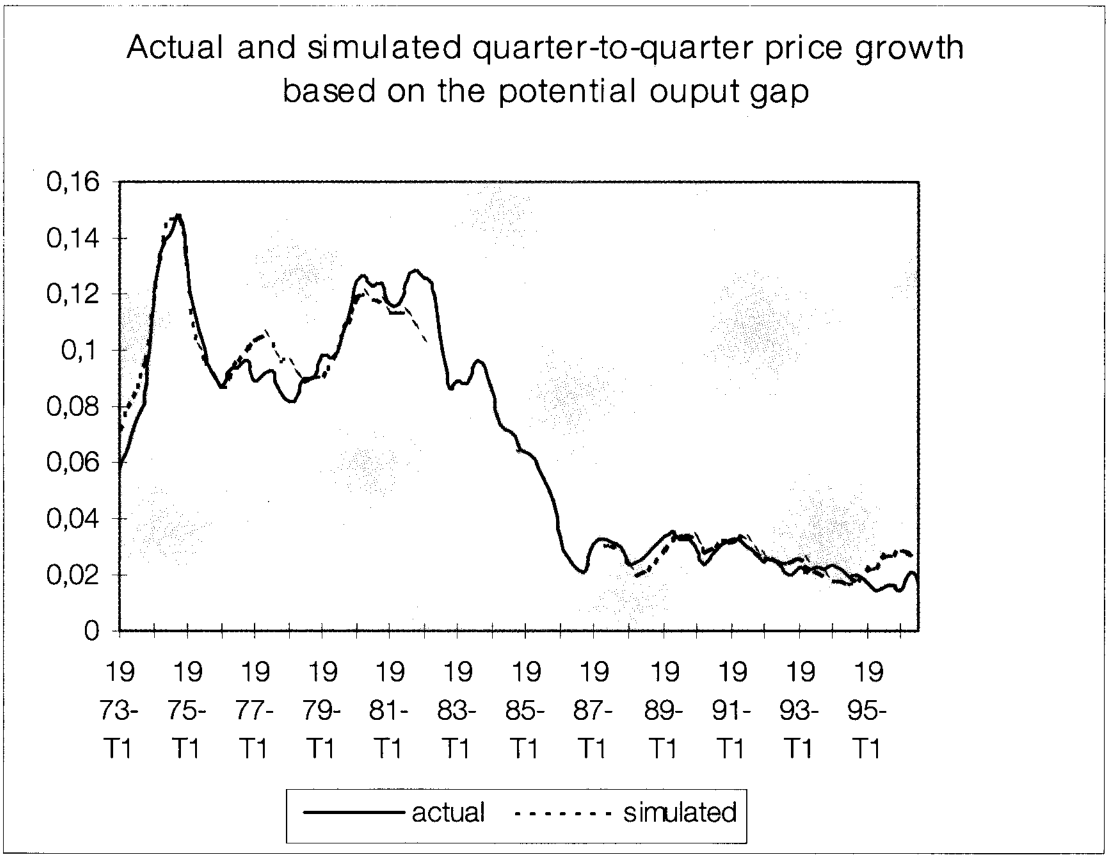
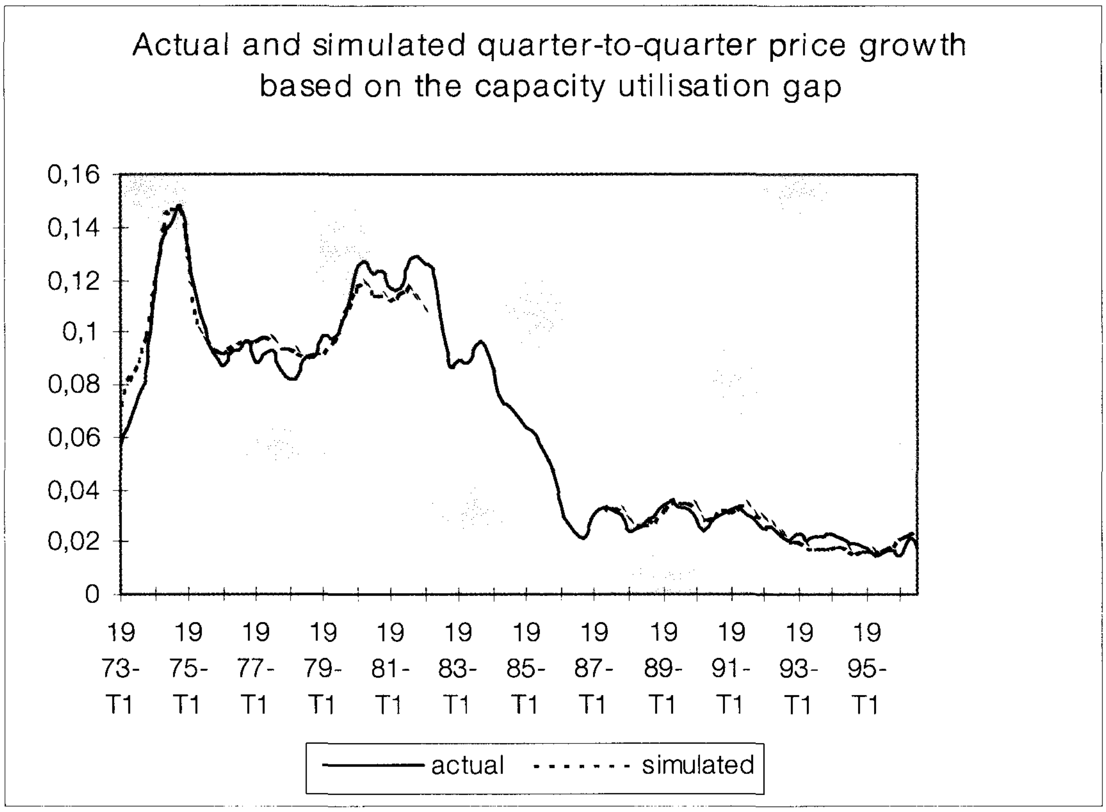


Chart 5



Conclusion

The methods proposed here for constructing output gap indicators and linking them to inflation meet three initial conditions : they are reproducible, easy to use and quick to produce results. They are based on a single or multiple-variable approach, including a structural component in the latter case. The results are mutually consistent and correspond to the results of the many other studies on the subject. The equilibrium rate of unemployment, defined here as the unemployment rate at which firms' marginal profit rates are stabilised in the short term, is calculated by comparing real pressures on wage growth with assumed pressures on the labour market, the latter being measured by the gap between the actual and equilibrium rate of unemployment. This method also makes it possible to identify the influence of changes in firms' net interest payments on short-term labour market pressures. These estimations, like all estimations of such indicators, are based on a large number of assumptions which we have endeavoured to clarify, though some of them are inevitably open to discussion. Our estimations should therefore be treated with the utmost caution.

The output or capacity utilisation gaps are found to exert an influence on price growth in France. The influence proves to be of the same size and does not appear to be affected by the inflationary climate. Whether inflation is high or low, there is an average inflation rate around which price growth fluctuates in response to internal and external pressures according to stable patterns of behaviour.

Annex 1

Firms' profit ratios and the equilibrium rate of unemployment: in search of a simple link

We shall first show, from a simplified wage-price loop and a few robust assumptions, how it is possible to link variations of the profit ratio with the gap between the actual and equilibrium rate of unemployment, and then go on to propose an empirical calibration of the parameter of this link.

A. Link between variations of the profit ratio and actual and equilibrium rates of unemployment

i) Profit ratio equation

$$TM = 1 - \frac{W.NE}{P.PIBE} \quad \Rightarrow \quad 1 - TM\frac{W.NE}{P.PIBE} \quad \Rightarrow \text{ by approximation:} \tag{A1}$$

$$\Delta TM = \dot{P} + \dot{\pi}_N - \dot{W} \text{ and} \tag{A2}$$

$$\dot{P} = \dot{W} - \dot{\pi}_N + \Delta TM \tag{A2'}$$

ii) Wage formation. Our starting point is a standard augmented Phillips-curve equation:

$$\dot{W} = \phi(L)\dot{P}_c - \beta TC + AV1, \text{ where: } \beta > 0, \text{ and } \phi(1) = 1 \text{ by assumption} \tag{A3}$$

$$\Rightarrow \quad \dot{W} = \dot{P} + \left(\phi(L)\dot{P}_c - \dot{P}\right) - \beta TC + AV1 \tag{A3'}$$

$AV = \phi(L)\dot{P}_c - \dot{P}$ represents the influence on wage growth of lags in wage indexation and deviations between $\dot{P}_c$ and $\dot{P}$ due to changes in the terms of trade or indirect taxation.

iii) TM is not directly influenced by the labour market*

$$TM = \frac{EBE}{P.PIBE} = \frac{EBE}{P.KE}.\frac{KE}{PIBE} \quad \Rightarrow \quad TM^* = \left(\frac{EBE}{P.KE}\right)^* . \left(\frac{KE}{PIBE}\right)^* \tag{A4}$$

$\left(\frac{EBE}{P.KE}\right)^* = TI + \delta$ if resources are well allocated and $\left(\frac{KE}{PIBE}\right)^*$ depends on technological factors.

The equilibrium profit ratio does not, therefore, depend on the labour market situation. ΔTM^* is assumed to be very weak, so that:

$$\Delta TM^* \approx 0 \tag{A5}$$

iv) From one period to another, changes in the equilibrium prove that:

$$\text{(A2'), (A3') and (A5)} \Rightarrow \begin{cases} \dot{P}^* = \dot{W}^* - \dot{\pi}_N \\ \dot{W}^* = \dot{P}^* + AV2^* - \beta TC^* + AV1 \end{cases} \Rightarrow \beta TC^* + \dot{\pi}_N - AV1 - AV2^* = 0 \quad \text{(A6)}$$

v) Changes in actual situations prove that:

$$\text{(A2') and (A3')} \Rightarrow \Delta TM = \beta TC^* + \dot{\pi}_N - AV1 - AV2 \quad \text{(A6')}$$

By differentiating (A6') and (A6), and assuming $AV2 \approx AV2^*$, we obtain:

$$\Delta TM \cong \beta\left(TC - TC^*\right) \qquad \textit{QED} \quad \text{(A7)}$$

B. Calculating the parameter β

From the previous equation we may assert:

$$\Delta^2 TM \cong \beta\left(\Delta TC - \Delta TC^*\right) \Rightarrow \beta \cong \frac{\Delta^2 TM}{\left(\Delta TC - \Delta TC^*\right)} \quad \text{(A7')}$$

Assuming $\Delta TC^* \cong \Delta TCl$ in the short term, we obtain:

$$\beta \cong \frac{\Delta^2 TM}{\left(\Delta TC - \Delta TCl\right)} \quad \text{(A8)}$$

The average value of the β coefficient obtained by applying equation (A8) is not homogeneous for the different profit ratios used and is even negative for the profit ratio at market prices (cf. Table A1, columns 1 and 3). This is because the denominator of the equation assumes values very close to zero in certain quarters or in certain years. We have decided to impose $\beta = 0.5$ as a standard value, for three reasons.

1. β = 0.5 is the value for which the various estimations of the output gap (using all the different methods described) are the same in 1993, the year in which the output gap was the smallest (approx. –2.0%) in the last ten years.

2. β = 0.5 corresponds to the average value resulting from application of equation (A8), for the four profit ratios under consideration and for the two smoothing parameters used for the HP filter (λ = 100 and λ = 1,600), if the quarterly values of β are reasonably bounded within the interval [0;1] (cf. Table A1, columns 2 and 4). These bounds seem reasonable because the β parameter cannot be negative. Moreover, if β = 0 the unemployment rate does not influence the profit ratio and β>1 the influence is negligible and there is little difference between the actual and equilibrium rate of unemployment (cf. Annex 2).

3. If equation (A6') is deduced from a standard wage-price loop, the β parameter corresponds to the effect of the level of unemployment on wages growth (equation (A3)). The calibration β = 0.5 corresponds to estimates of the augmented Phillips-curve equations used in the main French macroeconomic models (cf. G5M (1996)). In augmented Phillips-curve equations, when the level of unemployment influences wages growth (as in the Hermes model, for example), the

estimated β parameter is generally close to 0.5. When it is the logarithm of the unemployment rate that influences wages growth (as in the Amadeus model, for example), the estimated β parameter is close to 0.03 which, for unemployment rates in the region of 12%, is also consistent with a value β = 0.5.

Table A1

Average value of β obtained by applying equation (A8) to the interval 1970Q1 to 1995Q4

	Smoothing parameter of HP filter: λ = 1,600		**Smoothing parameter of HP filter: λ = 100**	
	Without bounds	**β bounded in the interval [0;1]**	**Without bounds**	**β bounded in the interval [0;1]**
*TM*1	3.14	0.48	10.72	0.50
*TM*2	-0.77	0.48	-2.83	0.50
*TM*3	1.41	0.49	2.88	0.51
*TM*4	1.19	0.51	2.38	0.53

Note: *TM*1: current profit ratio; *TM*2: profit ratio at market prices; *TM*3: profit ratio at factor cost; and *TM*4: profit ratio at factor cost excluding interest charges.

Annex 2

Table A2-1

Correlations between output gaps, gaps between actual and equilibrium unemployment rates and the production capacity utilisation rate (quarterly data: 1970-95)

	EPIBL	*EPIBT*	*EPIBP1*	*EPIBP2*	*EPIBP3*	*EPIBP4*	*TC1*–TC*	*TC2*–TC*	*TC3*–TC*	*TC4*–TC*	*TUA*
EPIBL	-	0.81	0.97	0.97	0.97	0.97	0.45	0.44	0.44	0.53	0.74
EPIBT	0.81	-	0.79	0.78	0.78	0.80	0.40	0.37	0.39	0.49	0.59
EPIBP1	0.97	0.79	-	0.99	0.99	0.99	0.58	0.57	0.57	0.63	0.68
EPIBP2	0.97	0.78	0.99	-	0.99	0.99	0.57	0.57	0.57	0.64	0.70
EPIBP3	0.97	0.78	0.99	0.99	-	0.99	0.59	0.58	0.59	0.65	0.69
EPIBP4	0.96	0.79	0.99	0.99	0.99	-	0.60	0.60	0.60	0.69	0.69
TC1–TC*	0.47	0.44	0.63	0.62	0.63	0.64	-	0.99	0.99	0.94	0.20
TC2–TC*	0.47	0.40	0.61	0.62	0.63	0.64	0.99	-	0.99	0.97	0.21
TC3–TC*	0.47	0.42	0.62	0.62	0.63	0.65	0.99	0.99	-	0.97	0.20
TC4–TC*	0.57	0.52	0.68	0.69	0.70	0.74	0.94	0.97	0.97	-	0.30
TUA	0.82	0.60	0.77	0.79	0.77	0.77	0.23	0.24	0.24	0.33	-

Table A2-2

Some results of econometric estimations explaining output gaps in terms of pressures in the labour and goods markets – explained variable: *EPIBPi*, with *i*: 1 to 4

Data	**Quarterly: 1970-2 to 1995-4**				**Annual: 1971 to 1995**			
	Explanatory variables			**R^2**	**Explanatory variables**			**R^2**
i	*TCi*–TC*	*TUA*	*Ct*	**DW**	*TCi*–TC*	*TUA*	*Ct*	**DW**
1	1.15		0.00040	0.33	1.18		0.00040	0.39
	(7.1)		(0.5)	0.47	(3.8)		(0.3)	1.63
		0.32	-0.27	0.47		0.34	-0.28	0.60
		(9.4)	(-9.5)	0.53		(5.8)	(-5.8)	1.06
	0.92	0.27	-0.23	0.68	0.90	0.29	-0.24	0.81
	(7.9)	(10.2)	(-10.2)	0.87	(4.9)	(6.9)	(-6.9)	2.03
2	1.05		0.00047	0.33	1.07		0.00048	0.38
	(7.0)		(0.5)	0.46	(3.8)		(0.3)	1.60
		0.33	-0.28	0.49		0.35	-0.29	0.62
		(9.8)	(-9.8)	0.54		(6.1)	(-6.1)	1.07
	0.82	0.28	-0.24	0.68	0.79	0.30	-0.25	0.81
	(7.7)	(10.5)	(-10.5)	0.86	(4.8)	(7.1)	(-7.1)	1.95
3	1.10		0.00038	0.34	1.12		0.00039	0.40
	(7.3)		(0.4)	0.47	(3.9)		(0.3)	1.62
		0.32	-0.27	0.47		0.34	-0.29	0.60
		(9.5)	(-9.5)	0.52		(5.9)	(-5.9)	1.04
	0.87	0.28	-0.23	0.68	0.84	0.29	-0.25	0.81
	(8.1)	(10.3)	(-10.3)	0.90	(5.0)	(7.0)	(-7.0)	1.98
4	1.14		0.00030	0.47	1.16		0.00031	0.54
	(9.5)		(0.4)	0.49	(5.2)		(0.2)	1.70
		0.35	-0.30	0.48		0.38	-0.32	0.59
		(9.6)	(-9.6)	0.46		(5.7)	(-5.7)	0.87
	0.87	0.27	-0.23	0.73	0.85	0.29	-0.24	0.85
	(9.6)	(9.8)	(-9.8)	0.86	(6.2)	(6.7)	(-6.7)	2.02

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Annex 3

Results of estimates of inflation equations

Table A3 - 1

Measurement of inflation with internal pressure indicators: external pressures based on import deflator growth

Alternative internal pressure indicators (*IPI*)	$\dot{P}_{c-1}$	IPI_{-2}	$\dot{P}_m$	**constant (x100)**	**break in constant (x100)**	R^2	**σ (x100)**
Smoothed ouput gap	0.29	0.09	0.09	1.57	-1.14	0.88	0.40
	(3.2)	(2.0)	(5.9)	(6.2)	(-6.9)		
Potential output gap	0.28	0.10	0.08	1.59	-1.14	0.88	0.40
	(3.0)	(2.2)	(5.9)	(6.3)	(-5.5)		
Capacity utilisation gap	0.33	0.03	0.09	1.51	-1.11	0.88	0.41
	(3.6)	(1.5)	(5.8)	(6.0)	(-5.3)		

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Table A3 - 2

Measurement of inflation with *smoothed*[12] internal pressure indicators: external pressures based on import deflator growth

Alternative smoothed internal pressure indicators (*IPIs*)	$\dot{P}_{c-1}$	$IPIs_{-2}$	$\dot{P}_m$	**constant (x100)**	**break in constant (x100)**	R^2	**σ (x100)**
Smoothed ouput gap	0.30	0.09	0.09	1.56	-1.13	0.88	0.40
	(3.2)	(2.0)	(6.0)	(6.2)	(-5.5)		
Potential output gap	0.29	0.11	0.09	1.55	-1.10	0.88	0.40
	(3.1)	(2.0)	(6.1)	(6.2)	(-5.4)		
Capacity utilisation gap	0.31	0.05	0.09	1.56	-1.15	0.88	0.40
	(3.4)	(2.1)	(6.0)	(6.3)	(-5.5)		

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Table A3 - 3

Measurement of inflation with internal pressure indicators: external pressures based on average foreign consumer price growth

Alternative indicators of internal tensions (*IPI*)	$\dot{P}_{c-1}$	IPI_{-2}	$\dot{P}_e$	**constant (x100)**	**break in constant (x100)**	R^2	**σ (x100)**
Smoothed ouput gap	0.29	0.12	0.07	1.68	-1.28	0.83	0.48
	(2.5)	(2.4)	(2.4)	(5.6)	(-5.2)		
Potential output gap	0.27	0.14	0.07	1.71	-1.27	0.84	0.48
	(2.3)	(2.5)	(2.4)	(5.7)	(-5.3)		
Capacity utilisation gap	0.33	0.05	0.07	1.61	-1.25	0.83	0.49
	(3.0)	(2.0)	(2.2)	(5.4)	(-5.0)		

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

---

12 $IPIs = (IPI + 1.5IPI_{-1} + 1.5IPI_{-2} + IPI_{-3})/5$.

Table A3 - 4

Measurement of inflation with *smoothed*[13] internal pressure indicators: external pressures based on average foreign consumer price growth

Alternative smoothed internal pressure indicators (*IPIs*)	$\dot{P}_{c-1}$	$IPIs_{-2}$	$\dot{P}_e$	**constant (x100)**	**break in constant (x100)**	R^2	σ **(x100)**
Smoothed ouput gap	0.30 (2.6)	0.11 (2.0)	0.07 (2.4)	1.65 (5.5)	-1.26 (-5.1)	0.83	0.49
Potential output gap	0.30 (2.6)	0.12 (1.9)	0.08 (2.4)	1.62 (5.4)	-1.21 (-5.0)	0.83	0.49
Capacity utilisation gap	0.31 (2.8)	0.06 (2.2)	0.07 (2.4)	1.64 (5.5)	-1.28 (-5.2)	0.83	0.48

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Table A3 - 5

Test of any break in the influence of the output or capacity gap on inflation from one sub-period to the other

Alternative indicators of internal tensions (*IPI*)	$\dot{P}_{c-1}$	IPI_{-2}	**break in the IPI influenc e**	$\dot{P}_m$	**constant (x100)**	**break in constant (x100)**	R^2	σ **(x100)**
Smoothed ouput gap	0.27 (2.8)	0.13 (1.9)	-0.07 (-0.8)	0.08 (5.6)	1.64 (6.2)	-1.19 (-5.5)	0.89	0.40
Potential output gap	0.26 (2.5)	0.15 (2.0)	-0.08 (-0.8)	0.08 (5.5)	1.66 (6.2)	-1.19 (-5.5)	0.89	0.40
Capacity utilisation gap (smoothed)	0.31 (3.2)	0.06 (1.9)	-0.03 (-0.6)	0.09 (6.0)	1.58 (6.3)	-1.16 (-5.5)	0.89	0.40

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Table A3 - 6

Test of any break in the influence of import deflator growth on inflation from one sub-period to the other

Alternative indicators of internal tensions (*IPI*)	$\dot{P}_{c-1}$	IPI_{-2}	$\dot{P}_m$	**break in $\dot{P}_m$ influenc e**	**constant (x100)**	**break in constant (x100)**	R^2	σ **(x100)**
Smoothed ouput gap	0.30 (3.1)	0.09 (2.0)	0.09 (5.6)	-0.01 (-0.1)	1.57 (6.1)	-1.14 (-5.3)	0.88	0.41
Potential output gap	0.28 (2.9)	0.10 (2.1)	0.09 (5.5)	-0.0 (-0.0)	1.59 (6.1)	-1.13 (-5.4)	0.88	0.41
capacity utilisation gap (smoothed)	0.31 (3.3)	0.04 (2.1)	0.09 (5.9)	-0.02 (-0.4)	1.55 (6.1)	-1.14 (-5.4)	0.88	0.41

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

[13] $IPIs = (IPI + 1.5IPI_{-1} + 1.5IPI_{-2} + IPI_{-3})/5$.

Table A3 -7

Measurement of inflation with alternative centred and reduced indicators of internal pressures

Alternative indicators of internal tensions (*IPI*)	$\dot{P}_{c-1}$	IPI_{-2}	$\dot{P}_m$	**constant (x100)**	**break in constant (x100)**	R^2	σ **(x100)**
Smoothed ouput gap	0.29 (3.2)	0.11 (2.0)	0.09 (5.9)	1.59 (6.2)	-1.14 (-5.5)	0.88	0.40
Potential output gap	0.28 (3.0)	0.11 (2.2)	0.08 (5.9)	1.61 (6.3)	-1.14 (-5.5)	0.88	0.40
Capacity utilisation gap (smoothed)	0.31 (3.4)	0.11 (2.1)	0.09 (6.0)	1.56 (6.3)	-1.15 (-5.5)	0.88	0.40

Note: The numbers in brackets correspond to the Student t-distribution of the estimated coefficients.

Bibliography

Allard, P. (1994): "Un repérage des cycles du PIB en France depuis l'après-guerre". *Economie et Prévision*, No. 112, 1994-1.

Banque de France, CEPREMAP, DP, ERASME, INSEE and OFCE (1996), (referred to in this paper as G5M (1996)): "Structures et propriétés de cinq modèles macroéconomiques français". *Mimeo*, *BdF-DEER Working Paper* No. 38, *INSEE-DEEE Working Paper* No. G9601.

Berger, K. and P. Teil (1996): "Calcul de la production potentielle en France et dans les grands pays de l'OCDE". *Mimeo*, SEMEF-BdF internship report, spring-summer 1996.

Boccara, F., B. Coeure and D. Eyssartier (1997): "Les économies dans le cycle". *Economie Internationale*, No. 69.

Bouthevillain, C. (1996): "Les cycles des grands pays industrialisés". *Economie et Statistique*, No. 298, 1996-8.

Bundesbank (1995): "Production potential in Germany and its determinants". *Deutsche Bundesbank Monthly Report*, August.

Cette, G. (1994): "L'efficacité du capital". *Economie Internationale*, No. 60, 4th quarter.

Cette, G. (1997): "Ecart de PIB et positionnement dans le cycle économique: quelques évaluations pour l'économie française". *Bulletin de la Banque de France*, No. 38, February.

Cette, G. and S. Mahfouz (1996): "Le partage primaire du revenu: un constat descriptif sur longue période". *Economie et Statistique*, No. 296-297, 6/7.

Cette, G. and D. Szpiro (1988): "L'appareil productif industriel: durée de vie des équipements, productivité et rentabilité". Banque de France, *Cahiers Economiques et Monétaires*, No. 28. The arguments relating to the lifetime and average age of capital goods were summarised in "Durée de vie et âge moyen de l'outil de production", *Economie et Statistique*, No. 208, March 1988.

Chatelain, J.-B., I. Odonnat and P. Sicsic (1996): "Construction sur longue période d'un indicateur d'inflation sous-jacente". Banque de France, DEER, *mimeo*.

Clark, P., D. Laxton and D. Rose (1996): "Asymmetry in the U.S. Output-Inflation Nexus". *IMF Staff Papers*, Vol. 43, No. 1, March.

Cour, P., H. Le Bihan and H. Sterdyniak (1997), (referred to in this paper as CLS (1996)): "La notion de croissance potentielle a-t-elle un sens?" *Economie Internationale*, No. 69.

De Masi, P. R. (1997): "Les estimations de PIB potentiel du FMI: Théorie et pratiques". *Economie Internationale*, No. 69.

Elmeskov, J. (1993): "High and persistent unemployment: assessment of the problem and its causes", *OECD Working Paper*, No. 132.

Elmeskov, J. and M. Macfarlan (1993): "Persisting unemployment". *OECD Economic Review*, No. 21, 1993/4.

Fisher, P. G., L. Mahadeva and J. D. Whitley (1997): "L'utilisation de mesures d'écarts de PIB pour prévoir l'inflation - l'expérience de la Banque d'Angleterre". *Economie Internationale*, No. 69.

Giorno, C., P. Richardson, D. Rosevaere and P. Van Den Noord (1995), (referred to in this paper as GRRV (1995)): "Potential output, output gaps and structural budget balances". *OECD Economic Review*, No. 24, 1995/1.

Giorno C., and W. Suyker (1997): "Les indicateurs de potentiel et d'écart de production de l'OCDE: Méthode d'estimation et utilisation pour l'analyse économique". *Economie Internationale*, No. 69.

INSEE (1995): "L'économie française - Edition 1995", Chapter on "Croissance potentielle, écarts de production et tensions sur les capacités de production". Hachette, *Livre de Poche*, Collection Références.

Jackman, R. and C. Leroy (1995): "Estimating the NAIRU: The Case of France". *Mimeo*, paper given at the EFTA conference, Lyon, 7-10th September.

Kahn, G. A. (1996): "New estimates of the U.S. Economy's potential growth rate". *Contemporary Economic Policy*, Vol. XIV, October.

Layard, R., S. Nickell and R. Jackman (1991): "*Unemployment*", Oxford University Press.

Maddison, A. (1993): "Standardised Estimates of Fixed Capital Stock: A Six Country Comparison". *Mimeo*, Essays on Innovation, Natural Resources and the International Economy, from *Innovazione e Materie Prime*, journal published three times yearly by the Ferruzzi Montedison Group, April.

Okun, A. (1962): "Potential GNP: its measurement and significance", in A. Okun (1970): "*The political economy of prosperity*", The Brookings Institution, Washington D.C.

Ongena, H. and W. Röger (1997): "Les méthodes d'évaluation des tendances et des écarts de PIB utilisées par les services de la Commission Européenne". *Economie Internationale*, No. 69.

Ponty, N. (1997): "Integration analysis of inflation rates in the G5 countries". *Mimeo*, BdF-Centre de Recherche.

Turner, D. (1995): "Speed limit and asymmetric inflation effects from the output gap in the major 7 economies". *OECD Economic Review*, No. 24, 1995/1.

Villetelle, J.-P. (1994): "Une fonction de production pour l'évaluation du PIB potentiel en France". *Mimeo*, BdF-SEMEF, No. 94-86, 30th October.

Westermann, T. (1997): "Le calcul des capacités de production utilisé à la Bundesbank". *Economie Internationale*, No. 69.

Comments on: "An empirical assessment of the link between the output gap and inflation in the French economy" by John Baude and Gilbert Cette

by Wilhelm Fritz

It is the purpose of this paper to analyse empirically the short-term relationship between variations in output during business cycles and changes in the inflation rate of the French economy.

In the first section the authors present three methods of calculating potential GDP: a Hodrick-Prescott filter, an estimation of trend-GDP which allows for breaks (the number and location of significant break-points being determined endogenously), and a third approach based on constant profit ratios. As it proves difficult to obtain a reliable estimate of the profit ratio defined as the coefficient on capital in a Cobb-Douglas production function, the authors set it equal to the average of the ratios between profits and GDP. No less than four such indicators are presented: two ratios are based on profits and GDP at, respectively, factor costs and market prices; a third relates profits at factor costs to GDP at market prices; and the fourth is a factor-costs-based profit ratio adjusted for net interest charges. One would assume that at least the first three definitions do not differ much in terms of their variability during the business cycle.

To derive unemployment and output gaps, for each of these profit ratios an "equilibrium unemployment rate" is determined such that no distributional pressures build up between capital and labour; i.e. the profit ratios have to be stable. Observed variations in profit ratios then determine the sign of the unemployment gap, and, by further assuming a proportional relationship, also its level. Changes in the unemployment gap determine, via the production function, the output gap, which obviously implies that the authors regard the size of the labour force as pre-determined. These relationships were not estimated econometrically.

This approach to determining the position in the business cycle is meant as an alternative to the derivation of a "non-accelerating wage or inflation rate of unemployment" (NAWRU or NAIRU). While the latter two can be interpreted as *structural* measures (i.e. the component of unemployment which is independent of the business cycle), the authors' equilibrium rates move with the actual rate during the cycle. They exceed the actual rate when wage growth accelerates and/or output price inflation decelerates. In these cases, employees are immediately penalised by less employment. In the definition where the profit ratio is calculated net of capital costs, employees also have to bear the burden of higher interest rates. Furthermore, even temporary changes in interest rates – such as the 1993 trough and the 1994 peak – feed immediately into the unemployment gap. Maybe for such reasons, the authors decided to smooth their profit ratio indicators to a certain degree by applying a Hodrick-Prescott filter before they calculated the unemployment gaps.

To conclude this section the authors present the various measures of the output- and unemployment gaps and the correlations between them. They also introduce a capacity utilisation indicator at this stage, arguing that this primarily captures pressures in the goods markets. Given the way that the unemployment gaps have been derived, it is hardly surprising that deviations of actual GDP and unemployment from their equilibrium levels are highly correlated. More to the point, it appears that equilibrium unemployment has increased in step with actual unemployment, so that their estimates are observationally equivalent to one that estimates equilibrium unemployment by a hysteresis model.

Section 2 examines the relationship between the position in the business cycle and inflation. To determine inflation, the authors use the constructed GDP gap indicators which captures, in one variable, inflationary pressures in both goods and labour markets. Alternatively, they replace the output by the capacity utilisation gap. More surprising in this context is their choice of a world-market price indicator instead of a variable which reflects inflationary expectations in domestic markets. The latter would have the advantage that it can be influenced by monetary policy while a central bank with a nominal exchange rate target has little influence on the former. The authors' decision to exclude the episode of a marked decrease in inflation rates between 1982 to 1986 from their sample has the advantage that they do not have to pay much attention to potentially non-stationary variables in specifying their reduced-form equations.

The overall result of these efforts is that increases both in capacity utilisation and actual output relative to potential significantly accelerate inflation and that domestic inflation responds almost equally strongly to changes in import prices. Given a high correlations between the capacity utilisation rate and the output gaps it is not surprising that the alternate specifications reveal the same dynamics; i.e. inflation responds to internal pressures with a two quarter lag. During the more recent sample period the simulated inflation rate exceeds the observed inflation rate, which raises the question whether the estimated relationships are actually stable beyond the sample period.

The output gap and inflation – experience at the Bank of England[1]

Paul G. Fisher, Lavan Mahadeva and John D. Whitley

Introduction

Modelling the UK economy at the Bank of England is based on the premise that no single methodological and empirical approach is likely to prove adequate to address the many economic issues which we typically face. Consequently a suite of models has been developed which span the range from highly theoretically representative agent models, based on calibrated data, to VAR-type models where economic theory plays a less well-defined role and statistical considerations are more dominant. This suite of models is described in Fisher and Whitley (1997). As part of this suite we have developed simple reduced-form models which summarise a more structural macroeconomic approach. An example of the reduced-form approach is the use of a measure of the output gap to summarise inflationary pressure as in a Phillips curve. This relationship can then provide a cross-check on forecasts of inflation arising from more detailed models. Its usefulness depends, however, on how well changes in the output gap predict changes in inflation. In this paper we describe our experience with this approach.

The first section of the paper sets out some of the methodological considerations. The second section describes alternative measures of the output gap and then derives a preferred measure together with an estimate of its uncertainty. The third section describes how measures of inflation expectations can be derived and this is followed by econometric evidence which relates the output gap and inflation expectations to actual inflation. Sections 5 and 6 look at issues associated with the Phillips curve: the possibility of sectoral bottlenecks; shifts in the short-run trade-off due to labour market flexibility; the possibility of "speed limit effects" and asymmetry in response to changes in the output gap. The final section sets out the conclusions.

1. Methodological considerations

The output gap is generally used to measure the extent to which the economy is operating at an unsustainable level of resource utilisation – often expressed empirically as the deviation of actual output from trend. In recent years it has usually been used to represent the extent to which the current level of output lies below the equilibrium level: sometimes called potential or full capacity output. Different implicit definitions of equilibrium partly explain different estimates. In ad hoc analyses the output gap is often measured as positive when output is below trend. When examining analytical models it is more consistent to define a positive value as output above trend. In what follows we always use the convention of a positive gap to mean that output is above trend.

In some alternative definitions – not used here – the output gap is defined in terms of how near the economy is to a maximum capacity ceiling; hence it is logically bounded at zero and always defined as positive. An example is full capacity based on peak-to-peak trend fitting of the business cycle.

---

[1] Preliminary results.

The output gap is used for two primary purposes – the analysis of inflationary pressure and cyclical adjustment of other variables, notably the public sector deficit. We concentrate here on the link to inflation.

1.1 An underlying theory of the output gap

In the event of a (positive) output gap caused by a positive demand shock, firms will employ more labour in the short-run for a given capital stock to produce the extra output to meet demand. To induce a greater supply of labour, firms must bid up the real wage rate (this is consistent with a large variety of labour market models). To work existing capital beyond its optimum can also bid up costs per unit of output. On both counts costs rise and, with a constant mark-up of prices over costs, prices also rise. Thereby real wages are reduced and the interaction of attempts to maintain real wages generates further rises in costs and prices (assuming that changes in the money stock accommodate higher demand). Inflation continues until policy reacts to offset the demand shock and reduce inflation if necessary. In the longer run the capital stock is unchanged and hence output should return to its previous level (or trend trajectory).

In the case of a (negative) output gap caused by a permanent and positive supply shock, higher productivity raises the return on capital, and should eventually lead firms to increase output. The extra level of output generates sufficient income to produce an equivalent level of demand. However, in the initial absence of higher demand firms will not expand output immediately. They thus need less labour, causing disinflationary pressure in a cycle opposite to that described for a demand shock. Disinflation only disappears if demand is increased and so, ultimately, policy must initiate, or at least accommodate, such an increase. In the longer run, the higher return on capital should lead to a natural expansion of output and demand through higher investment. In the case of a supply-side shock there is, therefore, a short-run and longer-run definition of potential in which the former is conditional on the existing capital stock. Hence there are corresponding different measures of the output gap. With the focus on inflationary pressure, we concentrate on the short-run definition.

An output gap can arise from either demand-side shocks or supply-side changes. In either case the output gap is consistent with disequilibrium in the labour market – sustained by an (dis)inflationary spiral – and sub-optimal capacity utilisation. Although the immediate consequence of an output gap is likely to be similar whatever the source, the longer-run implications are quite different as a supply-side shock can have a permanent effect on non-inflationary output levels (or even growth). To calculate the output gap one needs to be as clear about what is happening to potential as to actual output. Much of the debate on this topic is concerned with the measurement of potential, whereas the more simplistic calculations assume that the trend in potential output is either fixed or, at least, changes relatively slowly.

Although commonly used in analytical and relatively simple empirical models, the output gap rarely appears directly in structural macroeconometric models used for practical forecasting and simulation analysis. Such models typically focus directly on labour market pressure and, additionally, introduce direct pressure of demand in the goods market. For example, in the Bank's medium-term forecasting model, pressure in the goods market is captured by changes in the capital/output ratio. In the labour market there is a real wage bargaining model with a fixed natural rate of inactivity (variable by assumption). Deviations from the equilibrium inactivity rate and capital/output ratio together give the overall degree of inflationary pressure and hence the overall output gap.

The motivation for looking at the output gap is most simply illustrated by the Phillips curve. It is observed that fluctuations in activity are positively correlated with inflationary pressure. In what follows we illustrate this relationship using a generalisation of the expectations-augmented Phillips curve which can explain shifts in the observed relationship between inflation and activity. Note that in this general model, the *level* of inflation is not determinate unless supplemented by a nominal anchor through monetary policy. As such the Phillips curve only tells us about the dynamics of inflation. A further restriction on interpretation is that the Phillips curve makes no distinction between whether an output gap arrives through a demand or supply shock though the relationship is more likely to be robust for demand shocks.

The terms "output gap" and the "NAIRU" (non-accelerating inflation rate of unemployment) are often used interchangeably but are not identical. The NAIRU relates to equilibrium in the labour market whereas the output gap is a whole economy concept. For example, a shock to technical progress will change potential output, and hence the output gap, but will leave the NAIRU unchanged. In our discussion we concentrate on the output gap. In general we can write the relationship between the output gap and inflation as:

$$\Delta p = \phi \Delta p^e_{+1} + (1-\phi)\Delta p_{-1} + \gamma(y - y^*) \tag{1}$$

where Δp is the rate of inflation, Δp^e_{+1} is expected future inflation and y–y^* is the output gap. As the periodic re-setting of prices and wages is not synchronised across the different groups in the economy, the aggregate wage in part responds to expected future inflation and in part to its past value. The parameter ϕ thus measures how sticky inflation is. Expected future inflation is assumed to be given by:

$$\Delta p^e_{+1} = \mu \Delta p_{-1} + (1-\mu)\Delta p^* \tag{2}$$

where μ lies between zero and one and Δp^* is the government's inflation target. This means that individuals expect next period's inflation to be a weighted average of the government's target inflation and past inflation. The parameter μ measures how credible the government's target is: we have assumed that if μ is one, individuals believe that inflation will be unchanged from last period's value irrespective of the government's target. This gives the Phillips curve as:

$$\Delta p = \alpha + \beta \Delta p_{-1} + \gamma(y - y^*) \tag{3}$$

where $\alpha = \phi(1 - \mu)\Delta p^*$ and $\beta = \phi\mu + (1 - \phi)$.

In general α is non-zero while β lies between zero and unity, and the short-run Phillips curve is non-vertical. Remember that this assumes the existence of a monetary policy rule which delivers an average inflation rate of $\Delta p^* = \alpha/(1 - \beta)$. If the policy target – implicit or explicit – changes, then forward-looking rational agents will take account of this (the Lucas critique applies to this equation) and the parameter α will change. Ultimately, if policy has no credibility, μ and β tend to unity and α to zero such that inflation becomes indeterminate. This implies that a positive (or negative) output gap would cause inflation to be permanently increasing (or decreasing). If, on the other hand, policy is credible, inflation will move towards its new target at a rate depending on the degree of stickiness in wage and price setting.

Conditional on the parameter values and expectations formation, we have alternative policy scenarios for when a positive real demand shock hits the economy (for a negative shock the signs are simply reversed – in fact the first scenario assumes that negative and positive shocks are equally likely so as to maintain an average inflation rate). In the first scenario, we assume that policy is committed to an inflation target and economic agents fully believe in the authorities' commitment ($\mu = 0$) either explicitly or implicitly (e.g. through money growth rules); but inflation is sticky ($\phi<1$). The existence of an output gap puts upward pressure on inflation but the policy rule ensures that the gap is removed through a temporary policy tightening, so as to bring inflation back to target.

In the second scenario we also assume that the rate of inflation is sticky, while economic agents give a zero weight to the determination of the policy authorities to keep to an inflation target: $\mu = 1$, $\alpha = 0$ and $\beta = 1$. This implies that money growth is entirely endogenous such that there is no nominal anchor. In this case, inflation will rise unless the output gap disappears, so that bringing inflation back to the initial level requires an equal and opposite change in the gap – induced perhaps by a policy response.

A third scenario arises if inflation is perfectly flexible (ϕ=0) and an anti-inflationary policy is anticipated (μ=0). Inflation will now return to its initial level of its own accord as soon as the gap is closed.

The stylised facts for the United Kingdom prior to 1992 suggest a degree of price stickiness and less than full credibility. This corresponds to a case closer to the second than the first scenario. Low credibility may have occurred as the result of an insufficiently anti-inflationary policy. Inflation responds to a demand shock, increasing rapidly once agents realise that no policy action is forthcoming. When action is finally taken, policy not only has to compensate for the initial shock but must induce an equal and opposite shock to bring inflation back down again. The greater the degree of price stickiness the longer this process takes.

The monetary arrangements between September 1992 and May 1997 can be seen as approximating the first scenario; establishing credibility to help inflation to return to target, albeit slowly. Achieving the third scenario, with stable inflation and low output costs on the back of an anticipated and credible anti-inflationary policy, will be associated with even lower output costs if wages and prices are less sticky and such changes may themselves be facilitated by the more credible approach.

The value of the output gap approach in monitoring inflationary pressure (and thereby in setting policy) arises from the idea that a change in the output gap usually precedes the change in inflationary pressure it causes. Early action to counteract real shocks can thus minimise fluctuations in inflation. This lagged effect requires a degree of price stickiness whether or not this is accompanied by incomplete credibility.

The presence of lagged adjustment makes the dynamic processes much more complicated and introduces the possibility of "speed limit effects". Inflation may depend as much on the change in the output gap as the level. We discuss the relevance of these effects in Section 6 below.

An open economy affects the analysis in at least three ways:

(i) The direct effect of import prices means that there can be temporary external shocks to domestic retail price inflation. This can be expressed by adding a term to the Phillips curve for the deviation of the real exchange rate from equilibrium.

(ii) The existence of external trade means that capacity pressures may be offset by changes in imports/exports. However, in the absence of a change in relative prices, this tends to be a short-run effect since it is otherwise inconsistent with balance of payments equilibrium.

(iii) The external sector is an additional source of demand shocks.

2. Measuring the output gap

To measure the output gap one needs to estimate the unobservable level of potential output. The methods can be broadly summarised by the following categories:

(i) Smoothing and de-trending methods.

(ii) Econometric estimates (including production function estimates).

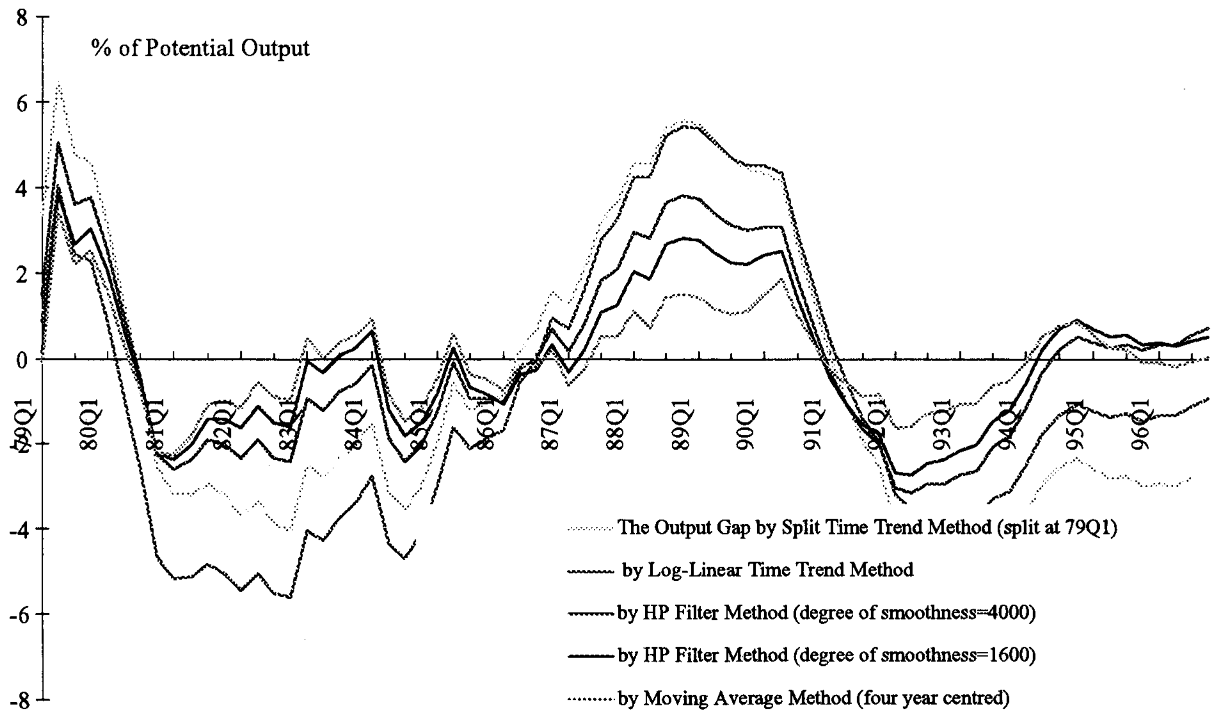
(iii) Survey data on capacity utilisation.

2.1 Smoothing and de-trending methods

The first category of methods are essentially statistical because they assume only that potential output follows some simple trend-like behaviour over time. They therefore focus on isolating this trend as a measure of potential output. Methods range from assuming that potential output grows at a constant exponential rate (the log-linear time trend); through assuming that it grows at a rate which is constant except for a few sudden breaks (the split time trend); to assuming that potential output can change freely but smoothly over time (i.e. the trend can "bend"). An example of a smoothing method is a moving average filter which calculates potential output at any point in time as a weighted average of the current future and past values of output around that point in time and thus evens out the cyclical effects (the filter should be of cycle length). Another popular example of a smoothing method is the Hodrick-Prescott filter (HP filter) which trades off the degree of smoothness in potential output against the extent to which it tracks actual output. This trade-off is determined by a parameter (λ, the degree of smoothness) which can be set to any value between the extremes of zero (where potential output is always equal to actual output and there is no output gap) and infinity (where potential output corresponds to the log-linear time trend). Although all de-trending methods are biased when a structural break has occurred, the Hodrick-Prescott method suffers from the additional disadvantage that it relies on extrapolations of data beyond the end of the period under consideration.

Chart 1

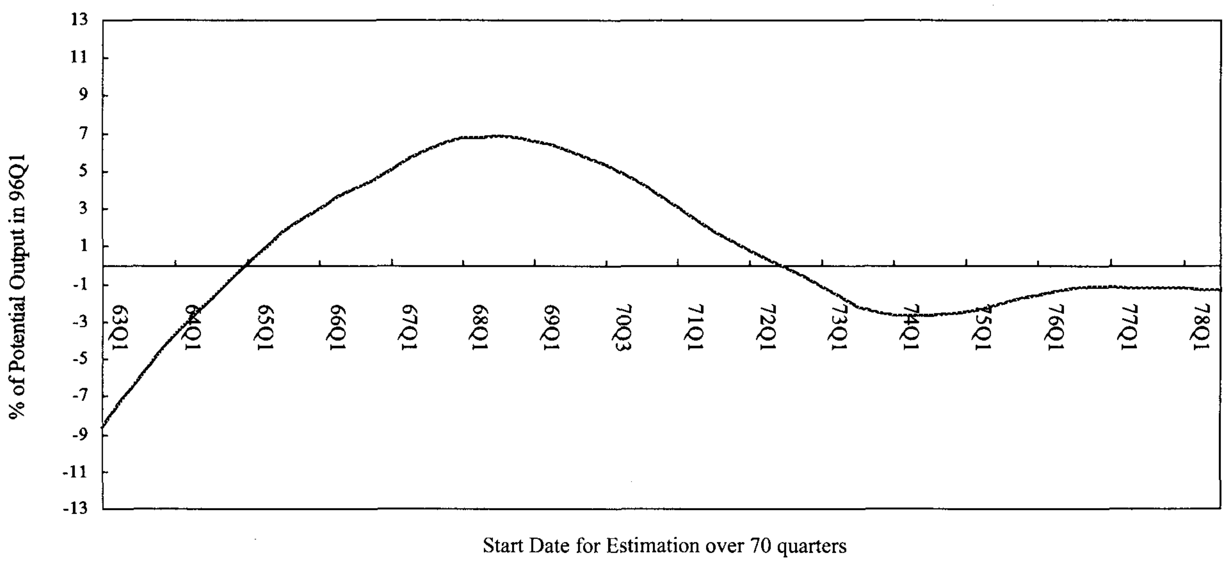
Output gaps by different smoothing and de-trending methods



In Chart 1 we present five measures of the output gap, corresponding to different trend methods. The HP filter is calculated for two degrees of smoothness, one of which (λ = 1,600) is that, typically found in the literature. Although there are similarities in the direction of change of the output gap over time, there is considerable variation in the most useful estimates: the value of the output gap at the end-point. This uncertainty is general as each method tends to produce a diverse range of estimates, depending on the particular assumptions made and the methods are always most uncertain about the current situation. For example, in Chart 2 we present the measure of the 1996Q1 output gap that would be given by using the log-linear time trend method to estimate a constant exponential growth rate for potential output using different time periods. Inclusion of the 1960s gives a significant positive gap in 1996Q1 but if, instead, we restrict our estimation to a sample beginning in the late 1970s, 1996Q1 would be a little less than 2% below trend. Further investigation using samples of the same size, over different periods, confirms the sensitivity of the 1996Q1 end-point estimate. This suggests that much of the disagreement about current estimates of the output gap in the United Kingdom should be cast in terms of economic arguments about what historical period is relevant for deriving the estimate.

Chart 2

The sensitivity of the 1996Q1 output gap to the start date by the log-linear time trend

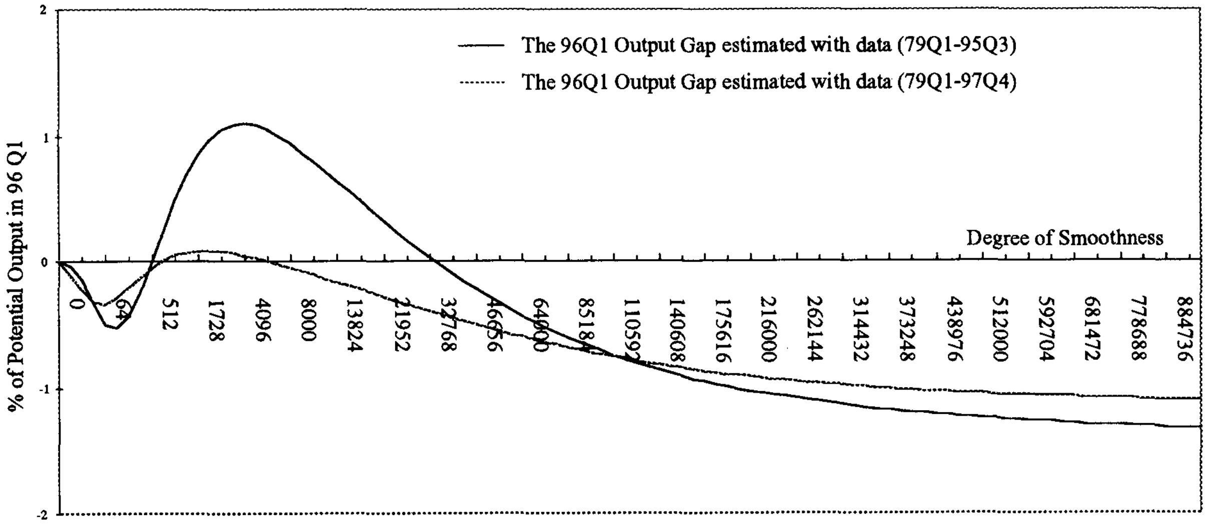


Similar problems are encountered with the HP filter method where the measure of the gap is sensitive to the degree of smoothing. Chart 3 presents the projections made with different values of λ with similarly dramatic results. The OECD has calculated the UK output gap using a value of 100 for λ and this would imply a small negative output gap for 1996Q1. A value of 1,600 gives a positive output gap of about 1%. As λ approaches infinity, the HP filter method gives the same result as would the assumption of a log-linear time trend for the sample used.

One plausible explanation as to why these smoothing and detrending methods fail to deliver robust results is that potential output has a random walk component rather than being a simple exponential or linear function of time. When this is the case, we would expect inherent instability in those detrending methods which fail to take account of this (Nelson and Plosser (1992) and Canova (1986)). This instability may manifest itself in the discovery of spurious output gaps (Nelson and Kang (1981)). There are measures of the output gap which begin by assuming that the potential output follows such a stochastic process (for example the Beveridge-Nelson decomposition). However, one problem with these methods is that there are an infinite number of ways in which a univariate time series can be decomposed into a trend and cycle (Quah (1992)). The issue here is that although each measure tends to give more stable estimates, there is a great divergence between the different methods and again, when a time series of GDP is considered in isolation, it is difficult to use economics to discriminate between different members of this class of estimates. The main conclusion is that there are no stylised facts about output gaps.

Chart 3

The sensitivity of the 1996Q1 output gap to the degree of smoothness in the HP filter method



2.2 Econometric methods

A natural response to this is to use economics to help identify differences in the measurement of potential output. Consider the following example using a simple Cobb-Douglas production function. Writing this in log-linear form (with lower case letters for logs):

$$y = \alpha n + (1-\alpha)k + mt + c \tag{4}$$

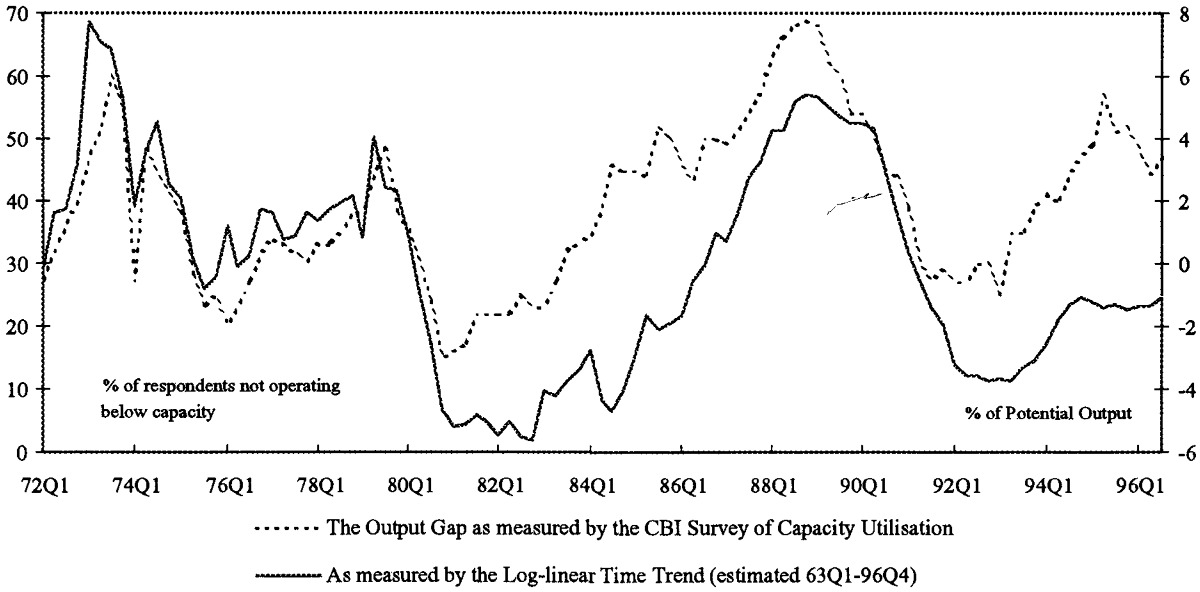
where y is output, n the labour force, k the capital stock and t a time trend with the parameter α giving the labour share. The values of the inputs n and k can be measured at equilibrium utilisation levels. Given an estimate of the labour share of income (α), we can estimate trend technological progress (m) by regressing the residual of log output over the estimates of the full capacity contribution of the factor inputs, on a constant and time trend. In our calculations, we have simply assumed that full capacity labour is given by the working population and capital by cumulated gross investment net of a constant depreciation rate.

Alternatively, econometric estimates of equilibrium output can be derived from some combination of theory and empirical estimation, such as the NAIRU (the non-accelerating inflation rate of unemployment). This involves using wage and price equations to determine directly the equilibrium rates of utilisation of labour and capital and comparing these with actual rates to determine a measure of the output gap. In principle, any uncertainty about the output gap measure can then be more easily linked to theoretical differences. The disadvantage is that these measures tend to be unstable over time. The Layard-Nickell approach (1991) is an example of how the NAIRU can be derived from wage and price behaviour. As noted above, the NAIRU and the output gap are not identical in concept and hence practice. We cannot, therefore, translate empirical estimates of one measure into another.

2.3 Survey evidence

The CBI Industrial Trends Survey asks firms whether they are operating at full capacity and gives a direct, though qualitative, indication of capacity utilisation (see Chart 4). Although this and other surveys tend to cover only the manufacturing sector, it can be generalised to the whole economy if other sectors have a common cycle with manufacturing. Problems can arise if the assumption about normal capacity levels changes over time in the survey responses. Furthermore, it is not clear whether firms take into account both labour and capital in assessing their level of capacity utilisation.

Chart 4
The output gap as measured by the CBI Survey of capacity utilisation in the manufacturing sector



2.4 Probability distribution of estimates of the output gap

Each method of measuring the output gap typically produces a diverse range of estimates, depending on the particular assumption used. It is not possible to discriminate between these measures in the absence of underlying theory. In consequence, we have greater faith in methods that are more theory-based, such as the production function approach. This incorporates views as to how labour and capital are combined to produce output.

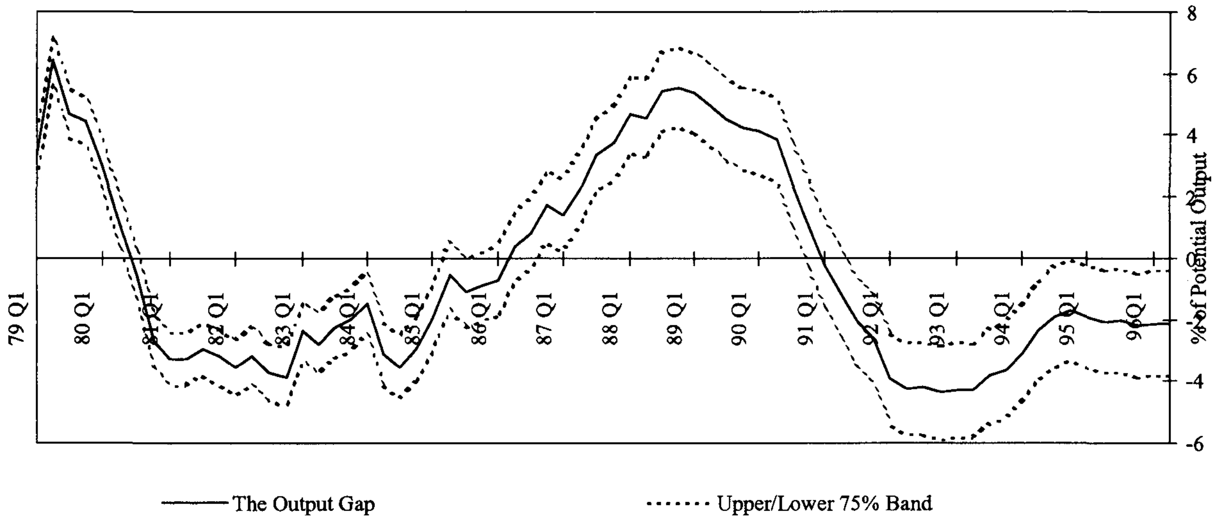
Furthermore, since point estimates of the output gap are unhelpful when there is a wide area of uncertainty, it is important to emphasise the degree of statistical confidence in the cases where we can derive the probability distribution of the estimate of the gap. Where no statistical measure of confidence can be derived there is an argument for not reporting estimates of the mean gap since it implies a misleading degree of precision. In the production function approach we can derive a confidence interval based on the econometric procedure used and in Chart 5 presents the bands of 75% certainty for an estimate using this methodology assuming a shift in technical progress after 1979.[2] The uncertainty in the estimate arises from the uncertainty in estimating the constant and the coefficient on the time trend in the equation described above. The precision of the estimate is also predicated on the assumption that we know the values of full capacity labour and capital, actual GDP,

2 The dating of this shift is confirmed by recursive estimates. Allowing for the shift implies that technological progress improves at an annual rate of 0.6% for the period beginning in the 1970s until 1979 and at a rate of 1.27% thereafter. Potential output as a whole grows at a faster (non-constant) rate because it depends on the rates of growth of equilibrium labour and capital as well as that of technological progress.

the labour share of income and the form of the production function with certainty. Relaxing these assumptions would widen the bands even further. The results show that the mean value of the output gap is -2.1% in 1996Q2 with a 75% confidence interval of -2.9% to -1.3%.

Chart 5

The output gap by the production function method



Our research has shown that the choice of time period is also critical to obtaining estimates from this method. Applying three different assumptions to quarterly real GDP data for the United Kingdom gives estimates of the output gap at the start of 1996 of alternatively 1.3%, -2.1% or -3.4%. The first estimate derives from the assumption that technical progress improves at a constant rate over the last 30 years (1963-93); the second that the whole process of technical progress has changed since 1979; and the third that there has been a change (increase) in the rate of improvement in technical progress after 1979, combined with a step loss in capacity in the 1980-81 recession. The rationale for a shift around 1980-81 is consistent with the effects of the shake-out of the manufacturing sector on aggregate productivity growth. We have adopted the second of these assumptions in deriving our preferred estimate of the probability distribution of the output gaps.

3. Measuring inflation expectations

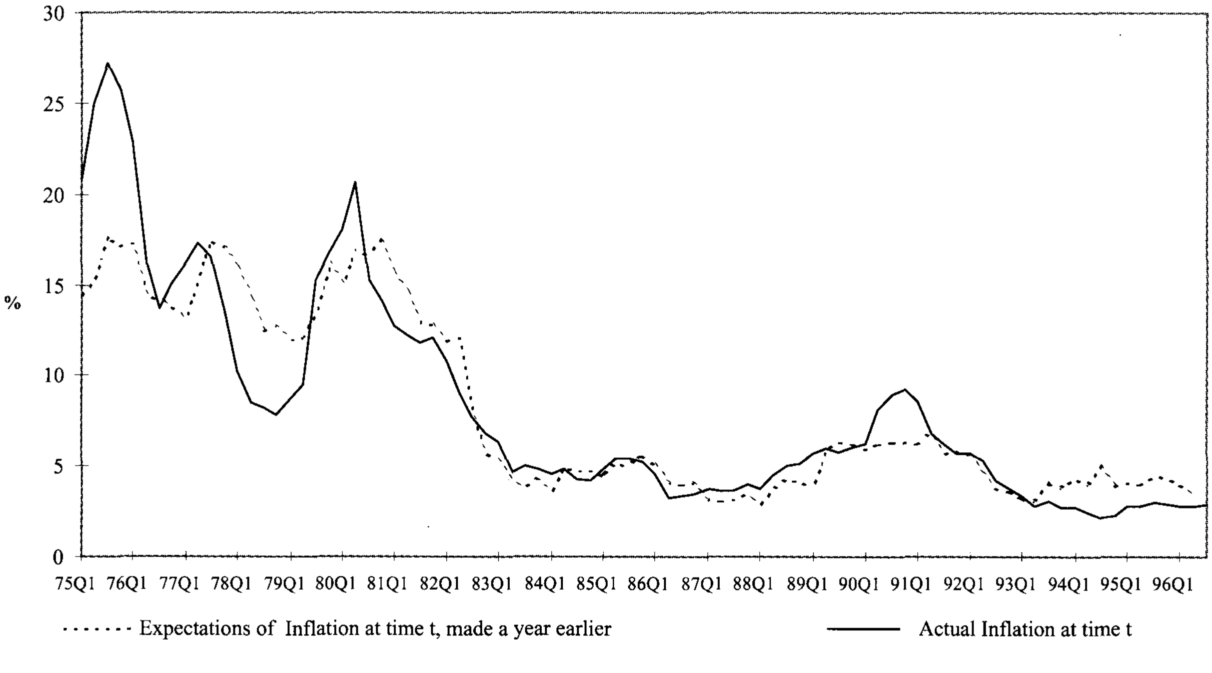
Inflation is affected by past values of inflation, economic agents' expectations of future inflation as well as the output gap. The coeffcient on the output gap captures the flexibility of prices: the more flexible are prices, the higher is the value of this coefficient. Substituting out for expectations of future inflation adaptively gives a formulation solely in terms of lagged inflation but such relationships may not be structurally stable. We have derived a direct measure of expected inflation to take account of potential instability from shifts in expectations when estimating the role of the output gap.

Expected inflation is constructed from the Gallup and GFK consumer confidence surveys. These surveys ask a sample of consumers to rank how strong they think inflation will be over the next twelve months. Responses are classified by whether they expect that, in relation to now, there will be a more rapid increase in prices; or that prices will increase at the same rate; or that prices will increase at a slower rate; or that prices will be stable or that prices will fall slightly; or that they don't know. The question on expected inflation has changed twice since the survey has been constructed; hence we construct the quantitative expectations measure over three sub-samples. We quantify the responses by summing the weighted proportions of the sample. The weights are simply determined by a time-series regression of actual inflation on the response proportions (see Pesaran (1988)). This implies that, on average, the weights are the same for actual inflation and expected inflation and the procedure, therefore, imposes unbiasedness on the expected inflation series. Hence tests of rationality are not possible.

The resultant series is shown in Chart 6 alongside actual (RPIX) inflation for the same period. One notable feature arising from comparison of actual and expected inflation rates is the unanticipated shocks due to the large oil price changes in 1974 and 1979. This suggests that the apparent breakdown of the Phillips curves in the 1970s was not entirely due to the lack of an explicit consideration of expectations but partly due to these unanticipated supply-side shocks. The chart also suggests that inflation expectations have behaved adaptively since 1987, lagging behind actual inflation. There does not appear to have been a shift in inflation expectations after 1993 on the introduction of the new monetary arrangements for the United Kingdom, which might have resulted in an increased credibility of economic policy. In particular, inflation expectations rose over 1994 and 1995 whilst actual inflation fell. This result is consistent with measures of inflation expectations derived from a comparison of indexed with conventional bonds. The failure of expectations to adjust to the new monetary regime may reflect the fact that it takes time to establish credibility together with concerns over the impact of sterling depreciation (both after ERM exit in 1992 and in 1995).

Chart 6

Expected and actual rates of inflation (annual rates of change)



4. Empirical estimates of the relationship between the output gap and inflation

We have used our preferred measure of the output gap derived from the production function approach, allowing for a shift in technological progress after 1979, together with our derived measure of inflation expectations to estimate a Phillips curve relationship. It is estimated on quarterly data for 1977Q1–1995Q1. The final form of the equation is derived by starting with a general form with several lags of the output gap and inflation and testing down until all the variables are significant. The final specification is:

$$\Delta_4 p_t = \phi \Delta_4 p_{t+4}^e + (1-\phi)\Delta_4 p_{t-1} + \gamma_1 \Delta\left(y_{t-1} - y_{t-1}^*\right) + \gamma_2 \left(y_{t-7} - y_{t-7}^*\right) + \left(\Delta_4 oil_t\right) \tag{5}$$

where $\Delta_4 p_t$ denotes annual RPIX inflation (logs), $\Delta_4 p_{t+4}^e$ expectations of inflation for period t+4 made at the beginning of period t, $y_t - y_t^*$ the output gap at time t, $\Delta\left(y_t - y_t^*\right)$ the change in the output gap from its last quarter's value and $\Delta_4 oil_t$ the rate of oil price inflation.

The underlying form of this equation is explained by New Keynesian theory. Nominal variables are sticky and the output gap is not eliminated in the short run. In addition, the presence of costs in adjusting prices implies that current inflation depends on previous values and expected future values of inflation. Different perceptions of policy, if credible, would be captured in the observed measure of inflation expectations. The oil price inflation term is included to capture errors on expectations made by the agents due to the oil price shocks of the 1970s. The parameters γ_1 and γ_2 measure the flexibility of price adjustment. If prices were more flexible then we would expect these coefficients to be larger; i.e. changes in the demand for output above full capacity lead to faster changes in prices.

The results of estimating this relationship over the period 1977Q1–1995Q1 are shown in Appendix A (Table A.1). The coefficients are all significant at the 5% level. The diagnostic tests reveal no evidence of specification errors apart from slight autocorrelation and non-normality. Autocorrelation in the residuals is increased if the estimation uses data from the early 1970s without the oil price inflation variable. The autocorrelation is thus due to the large temporary supply-side shocks which occurred in the United Kingdom at that time, a period which can be described as the graveyard of the UK Phillips curve. In part, this problem is resolved by including oil price inflation as an explanatory variable in the Phillips curve and in part by excluding data before 1977. Estimates of Phillips curves for the United States, which is a more closed economy and perhaps less subject to supply shocks, tend to find more robust reduced-form relationships (see, for example, Clark et al. (1995a)).

There is a fairly long lag on the output gap in the estimated equation; but, as we discuss later, interpretation of the dynamic response of inflation needs to take account of the endogeneity of lagged and expected inflation. The presence of non-normality in our errors raises the interesting possibility that the effect of the output gap on inflation is asymmetric, with a positive output gap exerting more inflationary pressure than the deflationary pressure exerted by a negative output gap of the same size. Equation errors may in fact be picking up this potential misspecification. We test for this possibility below and discover that allowing for asymmetries makes the residuals more normal.

As a further test of the specification we compare these results on quarterly data to a similar estimation performed on annual data reported in the Appendix (Table A.2). The estimation with annual data has less noise and hence a better fit but is also much less robust because the information content of annual data is too small. Although a comparison between different frequencies is not straightforward, there does seem to be some consistent features of the Phillips curve. In particular, the output gap affects inflation with a similar lag in both quarterly and annual data and, in general, the weights on backward-looking versus forward-looking inflation terms are similar.

The equation appears reasonably well-specified despite the considerable uncertainty as to any point estimate of the output gap. This is because it is the variation in the output gap estimate with respect to inflation that is important in least squares regression. As these different gap estimates move closely together, they presumably move with the true unobserved output gap. Therefore, the variation in the estimated output gaps used to calculate the coefficient in the Phillips curve is close enough to the true variation so as not to induce misspecification. To emphasise this further, we compare the fitted values to the values of actual inflation and expectations of inflation made one year earlier in the table below.

Out of sample forecasts of annual RPIX inflation

t=	95Q2	95Q3	95Q4	96Q1	96Q2	96Q3	96Q4	97Q1	97Q2
Expectations made at *t*-4	4.07	3.97	4.44	4.38	4.01	3.48	3.67	2.84	2.93
Forecasts	2.63	2.36	2.81	2.66	2.79	-	-	-	-
Actual inflation	2.75	2.93	2.92	2.83	2.79	2.87	-	-	-
Forecast errors	-0.12	-0.57	-0.11	-0.17	0	-	-	-	-

As confirmed by a Chow forecast test at a 1% level of significance, the out-of-sample forecasts are close to the actual values, even though point estimates of the output gap are uncertain (and agents make persistent errors in predicting inflation over this period).

4.1 Interpretation of the estimates

The estimation confirms that the effect of the output gap on inflation is unlikely to be instantaneous and may not just depend on the level of the output gap. The output gap begins to affect inflation after one quarter but this is temporary and subsequent effects appear with a further six-quarter lag. If the output gap is being closed quickly, "speed limit" effects create inflationary pressure more than a year before the level effect feeds through.

We can accept the restriction that the sum of the coefficients on expected and lagged inflation is unity at a 5% level of significance. This restriction means that inflation is dynamically homogenous. The coefficient on lagged inflation is about 0.8 and inflation is characterised as a sticky process.

These features together have an important implication for the inflationary process. If expectations are formed rationally, once we solve out for the expected inflation term, actual inflation is determined by past inflation with a unit coefficient (as well as the discounted path of past, current and future output gaps). This means that when agents are forward-looking, past inflation feeds one-to-one onto current inflation and there is an inherent tendency for the rate of inflation to be explosive. This does not mean that the rate of inflation will necessarily explode, or has no anchor, because the path of future output gaps determines where it will settle. Indeed, as the output gap eventually closes following a shock, the rate of inflation must eventually reach some constant equilibrium value which depends on the whole historical path of output gaps. It is the role of policy to bring inflation back to its long-run target by determining the current and future path of output gaps.

At first glance, it seems that inflation is not particularly sensitive to the level of the output gap with a 1 percentage point (pp) deviation in the gap from one quarter to the next generating a 0.55pp rise in the inflation rate after a quarter due to the "speed limit effect" and then only an additional 0.13pp after a year and a half. However, this interpretation fails to estimate the total effect because it considers only direct effects of the output gap on inflation.

More realistically, past and future inflation are also affected by the output gap and this feeds through to the path of inflation. In particular, the stickiness of inflation can make the long-term inflationary impact of a movement above potential much greater especially when policy is not credible in dampening future inflationary pressure.

To calculate the full effect of the output gap on inflation, we must determine the process by which expectations are formed. To illustrate this, we make two alternative assumptions and determine the inflationary effect of the output gap in each case. First, we assume that expectations are rational. We solve out for expectations of future inflation to find that actual inflation depends on the direct effect of the output gap, as before, as well as on the expected future path of output gaps and past inflation. Assume that the output gap increases by 1pp from an initial value of zero and then dies down gradually without any intervention from the policy maker. This can be interpreted as a simulation of an accommodated temporary increase in the output gap. The rate of inflation jumps immediately by 0.13pp because agents anticipate that the future output gaps are going to widen, and revise upwards their expectations of future inflation. The direct effect of the output gap raises inflation at one quarter, two quarters and seven quarter horizons and, over time, these direct effects reverberate on inflation, gradually disappearing as the gap closes by itself. The feedback of past inflation becomes increasingly dominant and, even after the output gap has closed, continues to drive inflation. Eventually, the rate of inflation settles 2.8pp higher than it was before the increase in the output gap. But this result depends on the speed at which the output gap is closed.

To illustrate the full effect of the output gap for an alternative expectations-generating process, we assume expected future inflation is as described in equation (6):

$$\Delta_4 p^e_{t+4} = \mu \Delta_4 p_{t-1} + (1-\mu)\Delta_4 p^* \qquad (6)$$

where $\Delta_4 p^*$ is the government's annual inflation target and the parameter μ measures how credible the government's target is such that the lower the value of μ the more credible is the target. We repeat the same exercise of increasing the output gap, letting it close by itself over time. Even with a small amount of credibility, the inflation rate eventually returns to its target value (2.5%) but because inflation is so sticky it can take a long time (typically ten years) to be within 0.1pp of its target. The speed of adjustment depends on credibility. With a degree of credibility corresponding to $\mu = 0.8$, the rate of inflation is 3.14pp after 3 years whereas with more credibility ($\mu = 0.8$) the inflation rate is 2.83pp at the same point.

The full effect of the output gap on inflation seems high when we take account of expectations, but it is worth emphasising that this scenario assumes that the policy maker does nothing to the future course of output gaps to bring inflation back to some target value.

5. Key issues

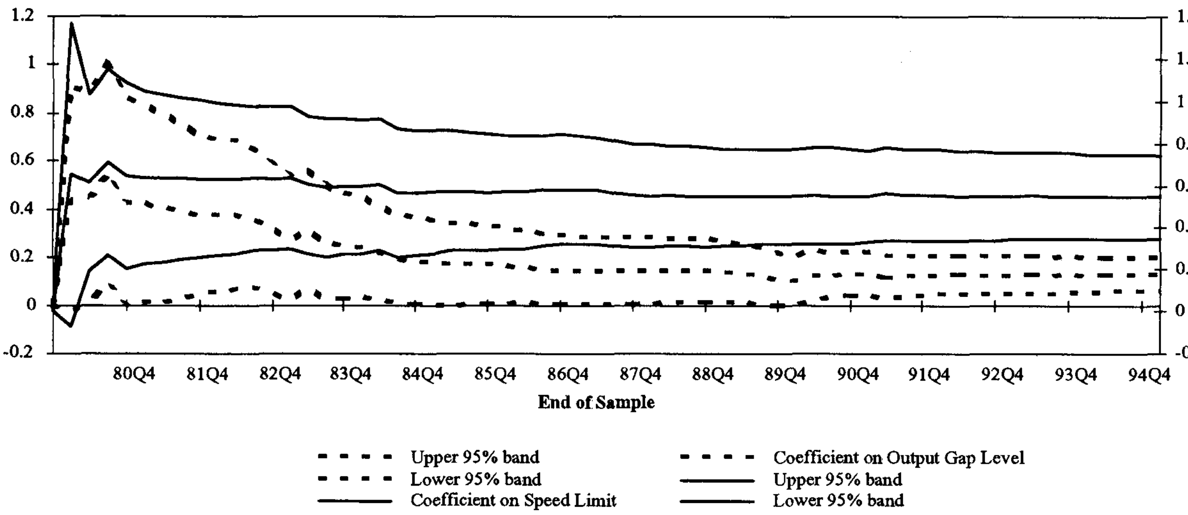
Once a stable relationship with inflation has been estimated, testing features of the framework with which short-term inflationary pressure is generated can be used to shed light on some key issues relevant to forecasting and policy. In this section we present two examples: one which arises from possible shifts in the trade-off between inflation and the output gap and the other which focuses on the bias in estimating inflationary pressure.

5.1 A change in flexibility in the labour markets in the 1990s

In this model, firms set prices as a fixed mark-up on wages and productivity. If workers become more willing to accept cuts in nominal wages as a result of legislation or market conditions, firms would be more free to change prices to absorb any change in demand. More flexible wages would in turn lead to a lower cost in changing prices and raise the parameters representing the trade-off between the output gap and inflation (γ_1 and γ_2).

We tested for a change in labour market flexibility by estimating the parameters γ_1 and γ_2 recursively over the sample. These are shown in Chart 7. Chow tests confirm the null hypothesis that there was no change in labour market flexibility represented by a shift occurring either jointly or separately in these parameters at any quarter from 1986Q1–1994Q4, even at very low levels of significance. It is important to bear in mind that more flexible wages do not necessarily arise from all changes towards more flexible labour markets. For example, lower hiring and firing costs would make it easier for firms to hold wages fixed by altering employment although it would also make it harder for unions to impose their targeted nominal wage.

Chart 7
Recursive estimates of coeeficients capturing flexibility



5.2 Testing for service sector bottlenecks

A further issue is whether sectoral considerations provide information about aggregate inflation, additional to the whole economy output gap. In particular we examine the possibility that the service sector may be more inflationary than other sectors. The aggregate Phillips curve may understate the degree of inflationary pressure because the service sector is more intensive in skilled labour. When the service sector is operating at a greater over-capacity than elsewhere it acts as a bottleneck generating extra inflationary pressure which would not be captured by an aggregate Phillips curve.

We can test this hypothesis with our aggregate Phillips curve estimates because it implies that the service sector output gap would have extra explanatory power in explaining inflation above the aggregate output gap and expected and past inflation terms. There is one complication to this straightforward test as the service sector output gap might affect inflation with a different lag than the rest of the economy. This is more in keeping with the idea of the service sector acting as a bottleneck and creating inflationary pressure before it arises elsewhere and even when there is significant downward pressure on inflation from other parts of the economy.

We deal with this by estimating a Phillips curve with several lags of both the aggregate output gap and the service sector output gap and testing down to see if the service sector output gap terms are significant. Allowing for a general lag structure takes account of the possibility of bottlenecks without imparting bias. The final specification, where none of the variables can be rejected at a 30% level, is given in the Appendix (Table A.3). A joint test on the significance of the service sector output gap terms rejects the hypothesis that they are insignificant and that there is no aggregation bias at even very low levels of significance. Although the diagnostic tests indicate some misspecification in the final estimates, our simple test is very supportive of the view that the service sector does act as a bottleneck. The results also indicate that the effect on inflation of allowing for differences in the service sector output gap can be sizeable. This can be illustrated with a simulation exercise. We compare the effects of a uniform rise in the output gap of 1pp everywhere with a 2pp rise in the service sector output gap accompanied by a lower rise or even a fall in the gap elsewhere. Both simulations imply that the output gap in aggregate increases by 1pp but the effect on inflation is quite different. Assuming that expectations are formed according to equation (6) with the degree of credibility given by $\mu = 0.4$, inflation is 3.5pp after three years when the rise is uniform and 4.7pp when the service sector acts as a bottleneck.

It is not easy to model these bottlenecks in an aggregate Phillips curve because they occur with a different lag and it is difficult to distinguish the extra inflationary impact of the service sector from the uniform tightening that occurs across all sectors. Estimating Phillips curves for each sector should help resolve this bias, but this requires assumptions about the relevant production functions and levels of factor inputs.

6. Speed limit effects and asymmetric effects in the Phillips curve

Short-term monetary policy would be easier if the effects of the output gap on inflation (given expected future and past inflation) were proportional even after a lag. Unfortunately, it may be more plausible to assume that the rate of closure of the gap itself (and whether it is positive or negative) also determines the amount of inflationary pressure a gap generates. In this section we examine how to measure these "speed limit" and asymmetric effects and whether they play a significant role in the UK economy.

6.1 Speed limit effects

It seems plausible that the rate of closure of the output gap can create inflationary pressure above that of the level of the output gap itself. It would be interesting to know how much difference this "speed limit effect" makes to the operation of monetary policy. If we can quantify the policy relevance of the "speed limit effect" then we can compare it across countries or scenarios. We would like to do this by measuring the extent to which the output gap cannot be closed because the "speed limit effect" evokes the threat of rising inflation. However, in a single equation this is not possible when inflation depends on past and expected future inflation.

The difficulty with measuring the "speed limit effect" as the coefficient on the change in the output gap in a single equation is that it depends on how we present the dynamics of the output gap and inflation. This can be explained by considering the following Phillips curve:

$$\Delta p = \phi \Delta p_{+1}^{e} + (1-\phi)\Delta p_{-1} + \gamma_1 \left(y_{-1} - y_{-1}^{*}\right) + \gamma_2 \Delta\left(y - y^{*}\right)$$

The same Phillips curve can be re-written as:

$$\Delta p = \phi \Delta p_{t+1}^{e} + (1-\phi)\Delta p_{-1} + (\gamma_2 - \gamma_1)\Delta(y - y^*) + \gamma_1 (y - y^*)$$

with a different coefficient on the change in the output gap. If we measure "the speed limit effect" as the coefficient on the change in the output gap we would find that the same Phillips Curve gives different measures of the speed limit effect (γ_2 and $\gamma_2 - \gamma_1$) depending only on how it is written.

The idea of a "speed limit effect" arises from a scenario where the output gap cannot be easily closed because of the threat of rising inflation. Thus Turner (1995) defines it with the following sentence: "If output begins by being below potential, then the greatest proportion of this gap that can be closed to keep next period's inflation from increasing is an inverse measure of the speed limit effect." If closing all the gap still keeps inflation constant then there is no "speed limit effect". If closing even a very small amount of the gap means that inflation starts to accelerate, then the 'speed limit effect' is strong.

Once we accept that current inflation is determined by its past and expected future values, this definition can be misleading. If inflation has just risen sharply (e.g. because of a supply shock), and is autocorrelated, it may be difficult, if not impossible, to stop it rising via demand management. Conversely, if inflation expectations are relatively constant then demand pressures on inflation will generally be much less. This measure of the speed limit effect which is based only on its threat to current inflation would report that the speed limit term was more important in the first case than the second even though the same Phillips curve underlies both scenarios. This shows that the intuition behind the "speed limit effect" can only be measured relative to where inflation is now; where it has been recently; and where it is expected to go.

The problem with the definition is that, if inflation is affected by past inflation and expected future inflation, it is more plausible that the policy maker would prefer a more gradual adjustment of inflation to its target value. These considerations make it clear that in order to make this policy experiment valid, we need to know the optimal path of the output gap over time so that we could examine the influence of the speed limit term on this path. In order to determine this optimal path we would have to consider not only the objectives of the policy maker but also other aspects of the transmission mechanism of monetary policy not captured by the Phillips curve equation.

We have explained why a measure of the policy relevance of a speed limit term depends crucially on the objectives and constraints of the policy-maker. This needs a fuller model of the economy and a historical context. We cannot adapt the single equation approach to adequately measure the "speed limit effect".

6.2 Asymmetric output gaps

It may be that the Phillips curve is asymmetric: a positive output gap exerts more inflationary pressure than the deflationary pressure exerted by a negative output gap of the same size. This can have very different implications, compared to a symmetric Phillips curve. For example, it implies that policy makers should err on the side of caution during a recovery because the costs of making mistakes and correcting past mistakes in monetary policy are higher than for a symmetric Phillips curve.

As Clark et al. (1995a) pointed out formally, a Phillips curve such as ours rules this possibility out for two reasons. First, the elasticities of positive and negative output gaps in affecting inflation are equal. Secondly, the OLS estimates of the output gap from past data place the same weight on positive as on negative deviations of actual output from estimated potential output implying that the economy has spent the same amount of time above and below potential. If, on the other hand, the true Phillips curve were asymmetric, output would have to spend more time in recession to make inflation (taking account of past and expected future inflation) stationary and the true output gap would be less than the calculated output gap $\left(y_t - y_t^*\right)$ by a constant α.

Incorporating both these considerations implies that we should estimate a Phillips curve of the following form:

$$\Delta_4 p_t = \phi \Delta_4 E_t p_{t+4} + (1-\phi)\Delta_4 p_{t-1} + \gamma_1 \Delta w_{t-1}\left(y_{t-1} - y_{t-1}^* - \alpha\right) + \gamma_2 w_{t-7}\left(y_{t-7} - y_{t-7}^* - \alpha\right) + \left(\Delta_4 oil_t\right) \quad (7)$$

where the weight w_t at each time t has to be estimated and itself depends on whether the output gap at this time $\left(y_{t-1} - y^*_{t-1} - \alpha\right)$ is positive or negative. The difficulty arises in estimating these weights along with the coefficient α. One simple way of doing this is to follow Turner (op cit) in assuming that the two forms of asymmetry (the weights and the coefficient) are related so that they can be estimated together. The estimated degree of asymmetry is that which best fits the data, and if this implies that the Phillips curve is symmetric, it should be preferred to a range of asymmetric alternatives.

Turner assumes that the weight on a negative output gap is a fixed proportion of the weight on a positive output gap of the same size, and calculates this proportion. We assume that the size as well as the sign of the output gap determine its weight. This exhibits the plausible property that inflation is stubborn and becomes increasingly harder to reduce when it is lower (and expectations of future inflation do not change) but easy to raise when it is high. In Appendix B we illustrate in more depth how this weight can be estimated and what it implies for inflation.

We found that the Phillips curve of best fit is not symmetric. Rather, we estimate that the true output gap is 0.55pp less than the calculated output gap. This means that when the true output gap is 2%, the inflation rate is 0.25pp more than the average rate; whereas when the output gap is -2%, the inflation rate is 0.20pp below the average rate. In addition we find that the asymmetric output gap terms are fairly significant in a regression with past inflation, expected inflation and symmetric output gap terms. The likelihood ratio test indicates that the asymmetric terms are significant at a 20% level. Although this latter test is biased against asymmetry, it supports our overall finding that there is some very mild asymmetry in the UK Phillips curve.

The results of re-estimating the Phillips curve with the asymmetric output gap terms are reported in the Appendix (Table A.4). The equation fits the data better than the symmetric Phillips curve as we should expect but without engendering any great changes in the other parameter estimates. There are still elements of skewness in the residuals but less than with the symmetric Phillips curve. These results do not necessarily contradict Turner who found that the Phillips curve for the United Kingdom was symmetric using annual data. This is because he was testing symmetry against the nearest alternative that a positive output gap exerted at least twice as much inflationary pressure as a negative output gap of the same size whereas our results estimate the degree of asymmetry to be too subtle to discern from his tests.

Conclusions

In this paper we have discussed the use of measures of the output gap and illustrated the diversity of estimates that can be obtained. We have shown how more theory-based estimates can be made and associated with a probability distribution for the output gap. We have further shown that, when coupled with an explicit measure of expected inflation, these output gap measures can be useful in estimating a relationship between the output gap and inflation. By considering the role of expected future inflation separately from the feedback of past inflation, we have a richer understanding of the dynamic effect of the output gap on inflation.

Using the estimated relationship between the output gap and inflation, we have shown that no change in labour market flexibility in the 1990s is apparent. There is some evidence, however, of the service sector acting as a bottleneck, generating more inflationary pressure than the rest of the economy. We have also discovered some mild asymmetry in the relationship between the output gap and inflation which implies that a 2% positive output gap exerts 0.05pp more inflationary pressure than the deflationary pressure exerted by a 2% negative output gap and that OLS estimates of the output gap are on average 0.55pp too high.

The simple idea of a "speed limit effect" asks whether the speed at which the output gap is closed has any bearing on the inflation rate; more strictly it asks what is the fastest rate of closure of the gap that can occur without an increase in the rate of inflation. Our discussion reveals that full appreciation of a "speed limit effect" needs to deal with the endogeneity of inflation expectations and the output gap itself. Inspection of the single equation Phillips curve is insufficient to answer this question. A proper conclusion requires a model of the output gap and the role of policy in influencing this gap, since the response of current inflation depends on past and future output gaps, and not only on the current level. Even without any bottlenecks or asymmetric effects, inflation may rise before any output gap is closed, or continue to change when output is growing at trend.

In conclusion, empirical estimation of expectations-augmented Phillips curves provides a role for the output gap in predicting inflation and the relationship can be robust when not faced with external or supply shocks. The output gap does, therefore, give us a simple method of judging short-term inflationary pressures alongside other modelling approaches. However, we emphasise that the single equation chosen has fairly limited use without some additional information about the determination of the output gap itself. The Phillips curve is not a reduced form in the proper sense of the term; it is just one part of a wider system.

Appendix A

Table A.1

An estimated UK Phillips curve equation on quarterly data[1]

	Annual RPIX inflation **Sample period: 1977Q1–1995Q1**
Output gap$_{t-7}$	0.279** (3.62)
Output gap$_{t-1}$ - output gap$_{t-2}$	0.5455** (5.42)
Expected inflation$_{t+4}$	0.217** (5.08)
Lagged inflation	0.783** (5.08)
Oil price inflation (spot price of Brent Crude)	0.0102** (3.61)
Equation diagnostics $\bar{R}^2$ DW SE Serial correlation (F) Heteroskedasticity (F) Functional form (F) Normality	 0.603 1.43 0.008604 LM(1) = 5.56 3.00 0.005 27.0**
Restriction that coefficients on expected and lagged inflation sum to unity	F=3.69*
Chow Forecast Test for 1995Q2–1996Q2 (F)	0.1

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.

[1] *t*-statistic in brackets.

Table A.2

An estimated UK Phillips curve equation on annual data[1]

	Annual RPIX inflation Sample period: 1977–1994
Constant	-0.00639 (1.2)
Output gap_{t-1}	0.1404* (1.8)
Expected $inflation_{t+1}$	0.3696** (7.3)
Lagged inflation	0.6362** (12.0)
Equation diagnostics	
$\overline{R}^2$	0.961
DW	2.33
SE	0.00898
Serial correlation (F)	LM(1) = 1.17
Heteroskedasticity (F)	0.85
Functional form (F)	0.002
Normality	9.0**
Restriction that coefficients on expected and lagged inflation sum to unity	F = 0.01

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.

[1] t-statistic in brackets.

Table A.3

An estimated UK Phillips curve equation on quarterly data allowing for service sector bottlenecks[1]

	Annual RPIX inflation Sample period: 1977Q1–1995Q1
Aggregate output gap terms at lags $_{t\text{-}1,\ t\text{-}2,\ t\text{-}4,\ t\text{-}5,\ t\text{-}7}$	0.263 (3.84), -0.441 (-3.58), -0.683 (-2.31), 0.146 (2.26), 0.120 (1.60)
Service sector output gap terms at lags $_{t\text{-}1,\ t\text{-}2,\ t\text{-}4,\ t\text{-}5}$	0.228 (0.837), -0.082 (-0.282), 1.03(3.72), -0.841 (-3.39)
Expected $inflation_{t+4}$	0.202 (4.52)
Lagged inflation	0.798 (4.52)
Oil price inflation (spot price of Brent Crude)	0.0103 (4.04)
Equation diagnostics	
R^2	0.695
DW	1.28
SE	0.007648
Serial correlation (F)	LM(1) = 19.42**
Heteroskedasticity (F)	1.375
Functional form (F)	0.513
Normality	12.0**

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.

[1] t-statistic in brackets (affected by multicolinearity despite estimation with reparameterisation).

Table A.4
An estimated asymmetric UK Phillips curve equation on quarterly data[1]

	Annual RPIX inflation Sample period: 1977Q1–1995Q1
Asymmetric output gap $_{t-7}$	0.111 (4.07)**
Asymmetric output gap $_{t-1}$ - asymmetric output gap $_{t-2}$	0.576 (7.20)**
Expected inflation $_{t+4}$	0.206 (5.63)**
Lagged inflation	0.794 (5.63)**
Oil price inflation (spot price of Brent Crude)	0.0104 (4.50)**
Equation diagnostics	
R^2	0.718
DW	1.83
SE	0.00687
Serial correlation (F)	LM(1) = 0.15
Heteroskedasticity (F)	2.78**
Functional form (F)	0.289
Normality	15.99**

* Indicates statistically significant at 10% level. ** Indicates statistically significant at 5% level.
[1] *t*-statistic in brackets.

Appendix B: Asymmetric effects

This technical appendix describes our method of estimating asymmetries in the relationship of output gaps to rates of inflation. This improves on past methods because it is less arbitrary and allows the size as well as the sign of the output gap to determine its effect on inflation. For ease of exposition, consider the simplest Phillips curve of the symmetric (linear) form:

$$\Delta p = c_0 + c_1(y - y^*) \tag{7}$$

As we described earlier, this differs from an asymmetric output gap in two ways. First, it imposes the same coefficient on positive and negative output gaps. Secondly, the OLS-estimated output gap ($y - y^*$) places the same elasticity on positive as on negative deviations of actual output from the estimated potential output, implying that the economy has spent the same amount of time above and below potential.

This second point should be considered in more depth. If the true Phillips curve were asymmetric and (quite reasonably) we assume that actual inflation is stationary,[3] then the economy

3 In our more general Phillips curve equation, we are assuming that actual inflation minus the effects of expected and lagged inflation is stationary.

would have spent more time below than above potential and the actual output gap in each period would be more negative than the calculated output gap ($y - y^*$) by a constant α.

Taking account of both of these aspects means that we can define the weighted output gap (*wtog*) as the calculated output gap adjusted for miscalculation and allowing for different weights for positive and negative values:

$$wtog_t = w_t\left(y - y^* - \alpha\right) \tag{8}$$

where w_t is the weight on the true output gap at time t and can itself depend on the output gap at time t. This weighted output gap will then have a linear (symmetric) effect on inflation.

Note that the sum of the weighted output gaps (just like the calculated output gaps) over the sample are zero; although the economy spends more time below trend, it affects inflation more when it is positive than when it is negative and the rate of inflation is stationary. This is true, even though the sum of the unweighted output gaps will be negative. Thus:

$$\sum_{t=0}^{n} w_t\left(y - y^* - \alpha\right) = 0 \tag{9}$$

Combining implies that, allowing for an asymmetry, we should estimate a Phillips curve of the following form:

$$\Delta p_t = c_0 + c_1 w_t\left(y - y^* - \alpha\right) \tag{10}$$

Remember that the weight at time t depends on whether the calculated output gap is greater than or less than α. The difficulty in estimating this Phillips curve is hence that the coefficient α and the weights should be estimated together.

Turner (op cit) suggested a simple test by assuming that the weight on a positive output gap is n times a negative output gap of the same absolute size. If the scale factor, n, is equal to one, then $\alpha = 0$ and the Phillips curve is linear (symmetric) as in equation 7.

For each value of the scale factor, we can calculate the value of α from equation 9 by a grid search. For this calculated α, the asymmetric output gaps are used in the estimation of 10. The fit of this estimation is compared with other estimations for different values of the scale factor. The scale factor which gives the best fit is chosen as best reflecting the DGP. If this optimal scale factor is one, then we have found that the symmetric Phillips curve is preferred to a range of alternatives.

There are two problems with this approach. First, for each scale factor, we have to find the optimal α by trial and error. Apart from being time-consuming, there may be more than one α which satisfies for any given scale factor, n. Secondly, as Turner himself notes, the form of the asymmetry is discrete whereas it may be that the weight an output gap receives in determining inflation depends on its absolute size as well as on its sign such that inflation becomes increasingly harder to reduce the lower it is.

Our alternative specification of the weights which overcomes both these problems is the exponential form:

$$w_t = \exp \rho\left(y - y^* - \alpha\right)$$

Here the more positive *and* the larger the output gap, the more weight it receives in determining inflation. The scale factor ρ plays an analogous role to the scale factor n in Turner's method: if it is zero then the Phillips curve is linear. Now, using equation (9), there is a unique, readily-calculable value of α for each ρ given by:

$$\alpha = \frac{\sum_{t=1}^{n}(y - y^*)e\rho(y - y^*)}{\sum_{t=1}^{n} e\rho(y - y^*)}$$

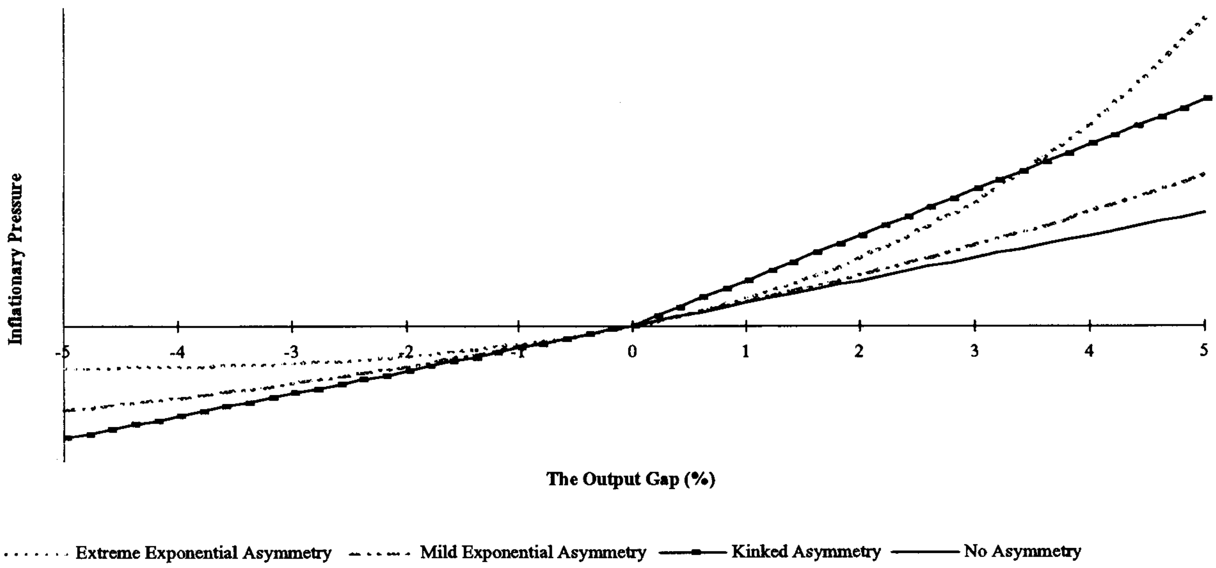
Thus for any value of ρ and given the data we can write down the value of α. This means that estimating the optimal degree of asymmetry is much more straightforward. The optimal degree chosen can then be used to derive policy implications.

Implications for inflation

Chart 8 compares the relative effect on inflation of an output gap according to the kinked weights used by Turner; our method for two degrees of asymmetry and a symmetric system of weights. The inflationary implications of each system differ. According to kinked weights, the elasticity of a given output gap only changes once it crosses full capacity. In contrast, with our weights, to reduce inflation to ever lower levels, ever deeper recessions are necessary.

Other specifications for the weights have been suggested in the literature. Clark et al. (1995b) illustrate the strengths and weaknesses of the different approaches. The system of weights they prefer is the method employed by Chadha, Masson and Meredith (1992) which assumes that the weights are given by a function which has similar properties to our exponential weights: a large negative output gap exerts relatively less deflationary pressure. However, for reasons similar to the kinked weights, this is not as easy to calculate as our system of weights.

Chart 8
Asymmetric output gaps



References

Batchelor, R.A. and A.B. Orr (1988): "Inflation Expectations Revisited". *Economica*, Vol. 33, pp. 317-31, August.

Beveridge, S. and C.R. Nelson (1981): "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to the Measurement of the Business Cycle". *Journal of Monetary Economics*, No. 7, pp. 151-74.

Canova, F. (1993): "Detrending and Business Cycle Facts". *CEPR Discussion Paper* No. 782, June.

Carlson, J.A. and M. Parkin (1975): "Inflation Expectations". *Economica*, Vol. 42, pp. 123-38, May.

Chadha, B., P. Masson and G. Meredith (1992): "Models of Inflation and the Costs of Disinflation". *IMF Staff Papers*, Vol. 31, pp. 395-431, June.

Clark, P., D. Laxton and D. Rose (1995a): "Asymmetry in the Inflation-Output Nexus: Issues and Evidence". *IMF Working Paper* No. 95/76.

Clark, P., D. Laxton and D. Rose (1995b): "Asymmetric Effects of Economic Activity on Inflation: Evidence and Policy Implications". *IMF Staff Papers*, Vol. 42, No. 2, pp. 344-74, June.

Dicks, M. (1996): "Output Gaps and All That". Lehman Brothers Global Economics, *International Economics*.

Fisher, P.G. and J.D. Whitley (1996): "Macroeconomic Model Projects at the Bank of England". Bank of England (forthcoming).

Giorno, C., P. Richardson, D. Roseveare and P. van den Noord (1995): "Potential Output, Output Gaps and Structural Budget Balances". *OECD Economic Studies* Vol. 1, No. 24, pp. 167-209.

Layard, R., S. Nickell and R. Jackman (1991): "*Unemployment: Macroeconomic Performance and the Labour Market*", OUP.

Nelson, C.R. and H. Kang (1991): "Spurious Periodicity in Appropriately Detrended Time Series". *Econometrica*, Vol. 49, No. 3, pp. 741-51, May.

Nelson, C.R. and C. Plosser (1982): "Trends and Random Walks in Macroeconomic Time Series". *Journal of Monetary Economics*, No. 10, pp. 139-62.

Pesaran, M.H. (1988): "*The Limits to Rational Expectations*", Basil Blackwell.

Quah, D. (1995): "The Relative Importance of Permanent and Transitory Components: Identification and Some Theoretical Bounds". *Econometrica,* Vol. 60, No. 1, pp. 107-18, January.

Roberts, J.M. (1995): "New Keynesian Economics and the Phillips Curve". *Journal of Money, Credit and Banking*, Vol. 27, No. 4, pp. 975-84, November.

Turner, D. (1995): "Speed Limit and Asymmetric Inflation Effects from the Output Gap in the Major Seven Economies". *OECD Economic Studies* Vol. 1, No 24, pp. 58-87.

Comments on: "The output gap and inflation – experience at the Bank of England" by Paul G. Fisher, Lavan Mahadeva and John D. Whitley

by Frank Smets

This paper nicely summarises some of the research done at the Bank of England on the relationship between inflation and the output gap. The paper consists of three sections. The first briefly discusses the concept of a Phillips curve and the various inputs needed for estimation, such as measures of the output gap and inflation expectations. The second presents estimates of a Phillips curve for the United Kingdom and discusses the implications of a widening output gap for future inflation. The third focuses on a number of related issues, in particular whether structural breaks can be detected, whether service sector bottlenecks have additional power, whether speed limit effects are important, and whether there are asymmetries in the link between inflation and the output gap. My comments will be limited to, first, estimation of the output gap and, second, some of the implications thereof for estimation of the Phillips curve.

Clearly, the usefulness of a Phillips curve for policy purposes critically depends on the reliability of the estimated output gap. The authors show quite convincingly that simple statistical methods, such as a log-linear trend or the HP filter, are very sensitive to the choice of the estimation period and the smoothing parameter in the HP filter. They acknowledge that there is a random walk component in output, but reject statistical models that take that explicitly into account on the argument that there are a million ways such a series can be decomposed into a permanent and transitory component. While this is true, it should be recognised that using economically sensible identifying assumptions and additional information can narrow down the set of possible decompositions. Examples of such fruitful approaches are the extended HP filter, the Blanchard-Quah methodology and the Unobservable Components model by Kuttner.

Instead the authors rely on a more structural approach which consists of using labour and capital inputs in a Cobb-Douglas production function and estimating the trend in productivity residuals using a constant and a linear time trend. The production function approach has its own problems related to the considerable data requirements, the possibility of introducing specification error because of, for instance, the assumed form of the production function and the difficulty in calculating confidence bands. More importantly, however, by fitting a linear time trend on the productivity residual, the authors run into the same kind of problems as the simple statistical methods they dismiss. It is a pity they do not report the diagnostics of the linear time trend regression, but I strongly suspect that the Durbin-Watson statistic is very low, indicating that the productivity residuals have a random walk component. If this is the case then not only will the regression on the time trend result in spurious correlation, it will also affect the calculation of the confidence bands. Basically one gets the same type of problems as in regressing output directly on a linear time trend. This is an old critique of the traditional production function approach (see, for instance, Schwert and Plosser (1979)).

The next question is then whether the estimation of the output gap matters for estimation of the Phillips curve. The authors find that it is mainly the change in the output gap that has a significant effect on inflation and, therefore, argue that it does not matter very much which output gap one uses, because they all move in a very similar way. My conjecture here is that this result may be a direct implication of the misspecification of the output gap. If there is still an important random walk component left in the estimated gap which has more to do with supply than demand conditions, then it should not be surprising that differencing helps in getting significant effects, because this implicitly filters out the low frequency or random walk component. To check whether this is what is going on, one should analyse the productivity residuals a bit further. Two pieces of evidence from Gerlach and Smets (1997) suggest that this is a possible explanation. Thus we find that the United Kingdom has the largest variance of random walk shocks among all the G7 countries. Second, once we filter out this quite variable potential output series, we find that the level of the output gap has a quite strong impact on inflation in the United Kingdom. Moreover, we tested for acceleration effects in our model but did not find any evidence of such effects.

References:

Gerlach, S. and F. Smets (1997): "Output gaps and inflation: evidence from the G7 countries". BIS *mimeo*.

Plosser, C. and G. Schwert (1979): "Potential output: its measurement and significance: a dissenting opinion". *Carnegie-Rochester Conference Series*.

Output gap and inflation: the case of Japan[1]

Tsutomu Watanabe

Introduction

The Japanese economy has been experiencing disinflation since the beginning of the 1990s when the "bubble" in stock and land prices burst. For example, the year-to-year inflation rate measured by the GDP deflator gradually declined from 3% at the beginning of 1992 to -1% at the second quarter of 1995. Moreover, if we call a fall in the price level deflation, the Japanese economy experienced deflation in 1994 and 1995. The inflation rate measured by the GDP deflator has started to go back above-zero since the third quarter of 1995 reflecting the recovery of the Japanese economy, but the trend of inflation is still very weak (see Figure 1 for the recent trend in various measures of inflation).

There are several discussions inside and outside the country about possible causes of disinflation and deflation. First, the real side of the economy has been very weak since the bubble burst, so that final demand for goods and services has also been weak. Second, reflecting the recent developments in the east Asian economies (i.e. the NIEs, ASEAN countries and China) and the appreciation of the yen during the period of 1990 to early 1995, cheap and good-quality products have flowed into the Japanese markets since 1993. A rising share of imported products as well as intense competition between imported and domestic products has significantly contributed to disinflation and deflation. Third, supply-side shocks within the country have also contributed to some extent. For example, the efficiency of the distribution system in the country has been improved significantly these five years: new types of large-scale shops called discounters or road-side shops have emerged and sell imported and domestic products at very competitive prices; the notoriously inefficient hierarchy of wholesalers has collapsed, at least in part. Also, deregulation in areas such as gas stations and automobile inspection has urged competition in those industries, thereby contributing to reducing their prices.

Common to the above stories is that "excess supply", whatever it might mean, has played an important role in the process of disinflation and deflation. In this paper, based on this understanding, we will first estimate the GDP gap as a measure of "excess supply or demand" and then investigate the relationship between the GDP gap and inflation.

The paper is structured as follows. In the next section, we will estimate the GDP gap. In Section 2 we will investigate the relationship between the GDP gap and inflation, in particular, whether (i) we can observe a stable Phillips curve; (ii) a concept like NAIRU could be applied to the Japanese economy; and (iii) the so-called "speed limit effect" as discussed in Romer (1996) and others, can be observed. Section 3 will conclude the paper by briefly considering the cost of disinflation with a special attention on the recent discussions about the cost and benefit of zero inflation.

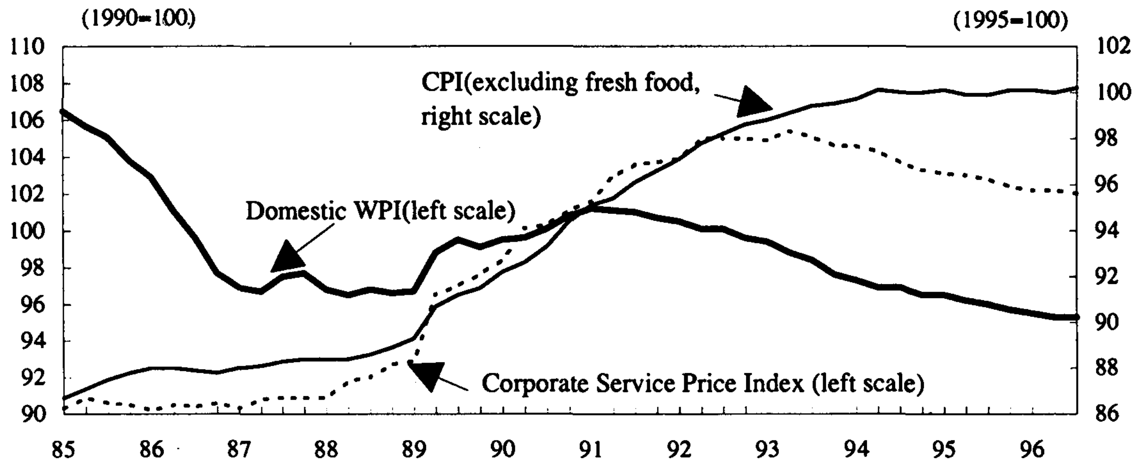
---

1 Please do not quote without permission. Part of the analysis in this paper is based on the products of other research projects in the Research and Statistics Department of the Bank of Japan. I want to thank members of those projects for their cooperation. Also, I would like to thank participants at the Central Bank Econometricians' meeting held at the BIS, particularly Steve Kamin, for their helpful comments and suggestions. Needless to say, however, all remaining errors are mine.

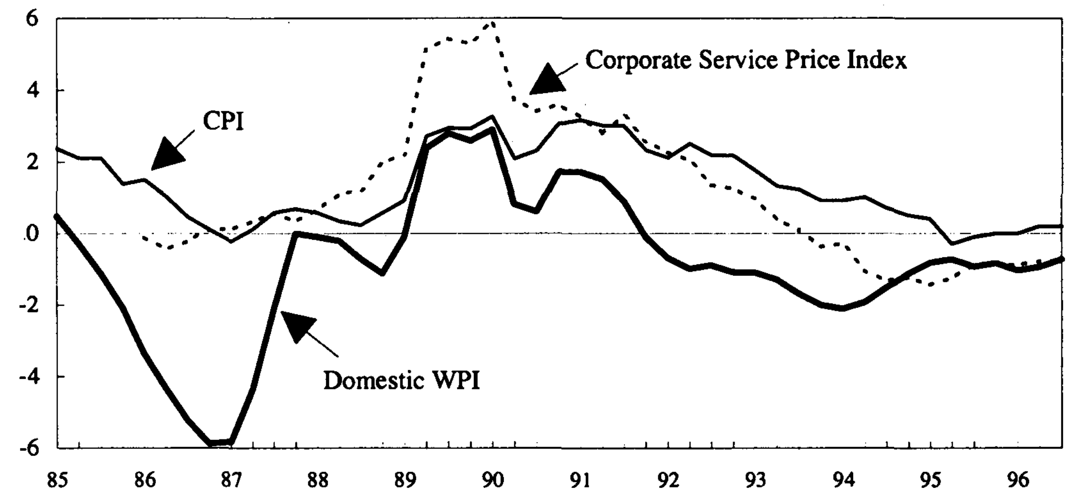
Figure 1

Recent trend in indicators of inflation

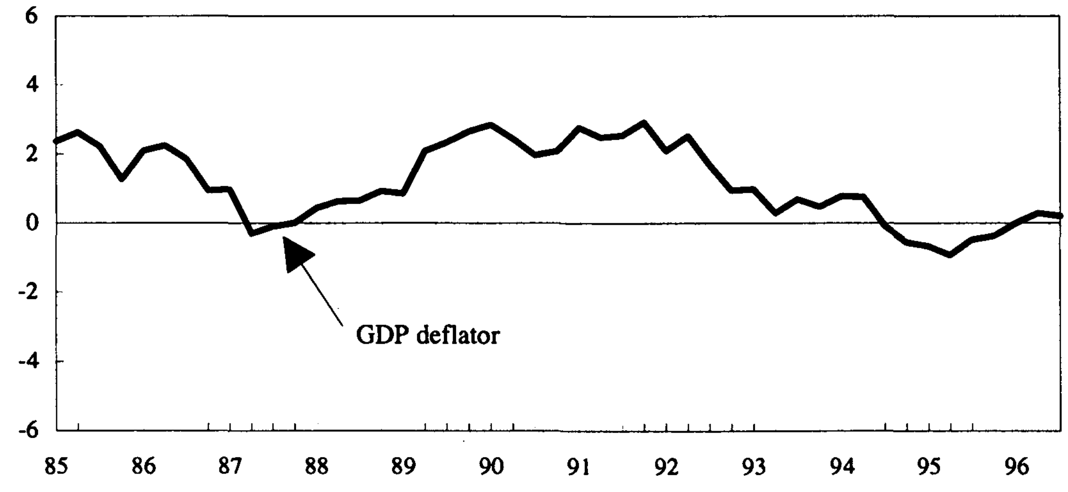
1. Levels of price indexes



2. Year-to-year change



3. Year-to-year change



1. Estimation of the GDP gap

1.1 Production function

To estimate the GDP gap, we first estimate potential output and then calculate the gap as the deviation of actual GDP from potential output. The procedure of estimating potential output is in three steps: (i) specify and estimate a production function of the Cobb-Douglas type; (ii) estimate the potential amounts of labour as well as capital inputs; (iii) put the potential amounts of factors in (ii) into the production function estimated in (i) to obtain potential output.

To begin with, let us assume the following production function:

$$\ln (Y_t / L_t H_t) = a_0 + a_1 \ln (K_t / L_t H_t) + a_2 t_1 + a_3 t_2 \qquad (1)$$

where Y, L, H, and K represent real GDP, employed persons, total working hours and capital stock multiplied by the rate of capital utilisation, respectively. Equation (1) is standard except for two things. First, since no data are available for the rate of capital utilisation in the non-manufacturing sector, we assume that the capital utilisation in that sector is constant. More precisely, we define K as $K = K_m R + K_n$, where K_m and K_n represent the capital stock in, respectively, the manufacturing sector and the non-manufacturing sector and R is the capital utilisation rate index in the manufacturing sector. Second, the way we express technological progress is a bit unusual. That is, we believe that the growth rate of total factor productivity, or TFP, is the same before and after the bubble period, but takes a different value during the bubble period. To express this idea, we introduce two linear time trends, the third and fourth terms on the right hand side of equation (1): the third term is a linear time trend for the entire estimation period (1975 to 1996); the fourth term is a linear time trend only for the bubble period (1985 to 1993). By construction, the growth rate of TFP is a_2 for the normal period and a_2+a_3 for the bubble period.

When equation (1) is estimated by OLS over 1975Q1 to 1996Q3, the result is:

$$a_0 = -2.421; \quad a_1 = 0.320; \quad a_2 = 0.0017; \quad a_3 = 0.0027 \qquad (2)$$

All of the parameters are different from zero at the 1% significance level. The estimated values for a_2 and a_3 mean that the TFP grows at 0.68 (= 0.17 x 4) % per year before and after the bubble period and by 1.76 (= 0.17 x 4 + 0.27 x 4) % per year during the bubble period.

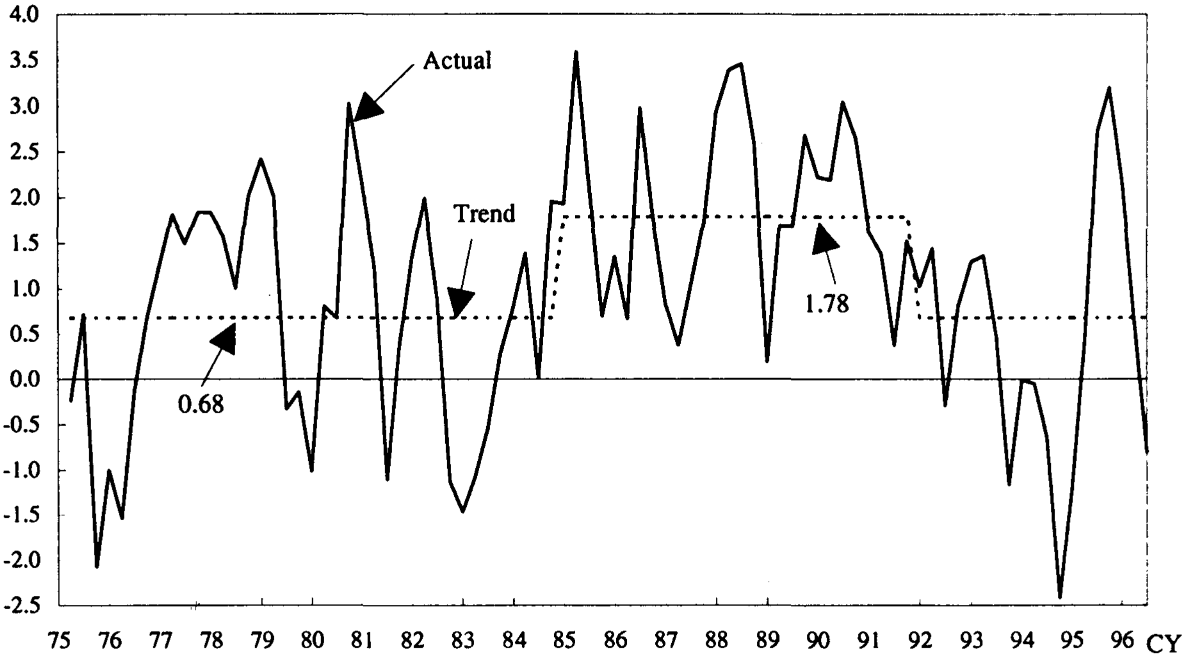
We should make one comment concerning the above result. We can calculate "actual" values of TFP as a residual: $\ln (Y_t / L_t H_t) - a_0 - a_1 \ln (K_t / L_t H_t)$. The solid line in Figure 2 is the growth rate of actual TFP calculated in this way. We expect that the growth rate of actual TFP is near a_2 for normal periods and a_2+a_3 for the bubble period. Indeed, this is true during the pre-bubble period as well as the bubble period. However, the actual growth rate of TFP is much lower than a_2 after the crash of the bubble. If the growth rate of TFP has actually fallen during the post-bubble period, we would need to add another linear time trend for that period. But the problem here is that, as seen in Figure 2, the actual growth rate of TFP was negative in 1994 and 1995. If we added a time trend for the post-bubble period, we would get a negative coefficient. This means that technological *retrogress* occurred in this period and, as a result, potential output grew less than the increase in factor inputs. Clearly this is not plausible and seems to imply that something is wrong with the estimation procedure. For example, effective labour inputs might be much smaller than the actual number of employed persons multiplied by working hours: in other words, the unemployment rate inside the firms might be higher during the post-bubble period. Or, the production function might have increasing returns to scale rather than constant returns to scale as we assumed in the estimation. If this is the case, a significant decrease in demand leads to larger decline in output than in factor inputs.

Although we can think of several reasons, we cannot give a definite answer at this moment. Given the current state of understanding, we do not believe that adding a time trend for the post-bubble period is the best strategy. Recognising that this is still an open question, we take the position in the rest of this paper that the trend growth rate of TFP in the post-bubble period is the same as the pre-bubble period.

Figure 2

Total factor productivity

Percentage change from previous period at annual rate



1.2 Potential inputs

Next, we estimate potential inputs. We define potential capital inputs as the actual capital stock multiplied by the historically maximum value of the capital utilisation rate in the manufacturing sector during the sample period (100.9 in 1990Q4 (see Figure 3)) and use this value throughout the sample. As for the non-manufacturing sector, we ignore capital utilisation as stated before.

We define potential labour inputs as the maximum number of workers multiplied by the maximum number of working hours. To estimate the maximum number of workers, we first calculate the ratio of employed persons to the population aged 15 years and over; i.e. the labour participation rate for the generation of 15-65 years as well as that for the generation of over 65 years, and then estimate trends in labour participation rates in each generation (see Figure 4). Multiplying the trend obtained in this way by the population, we get the maximum number of workers.

The maximum number of the overtime working hours in the sample period was 15.8 hours per month in 1989Q2. We use this value throughout the sample. As for scheduled working hours, there is a downward trend starting in 1988 and ending in 1993 which reflects the fact that more and more firms were adopting a five-day working week during this period (see Figure 5). Taking this observation into consideration, the potential of scheduled hours is estimated to be 162.2 hours per month in and before 1987, decreasing in 1988-93 as depicted in Figure 5, and to 148.6 hours in and after 1994.

Figure 3

Capital utilisation in the manufacturing sector

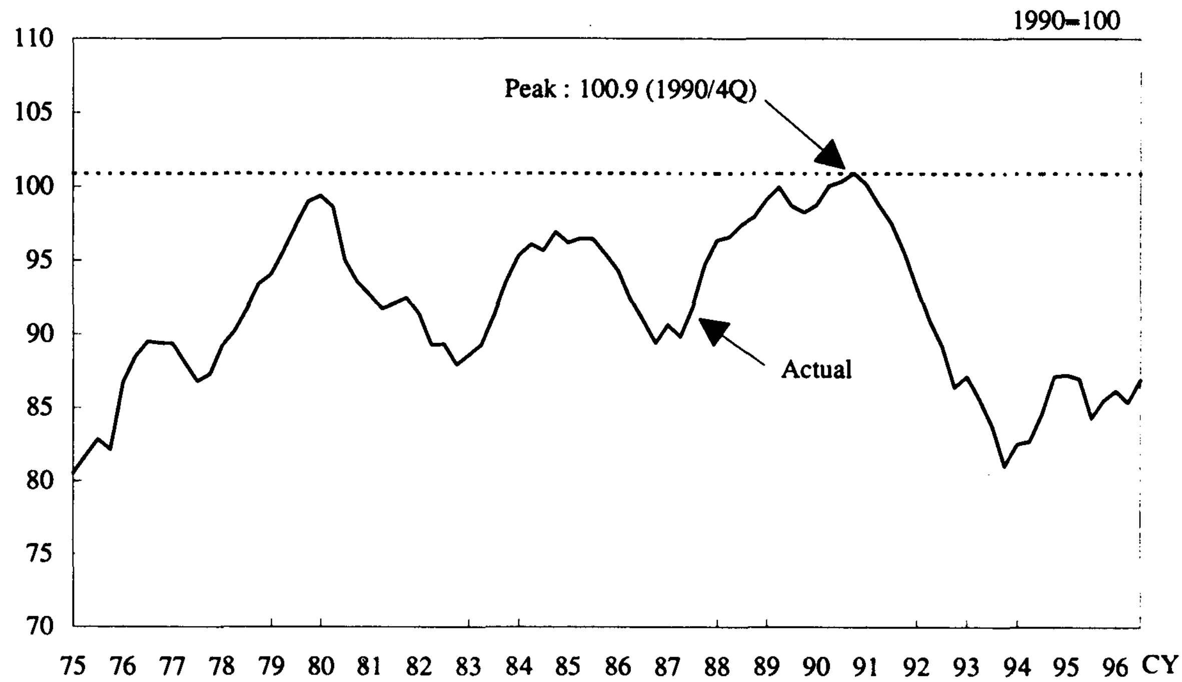


Figure 4

Labour participation rate

Percentage of employed persons in the population of 15 years old and over

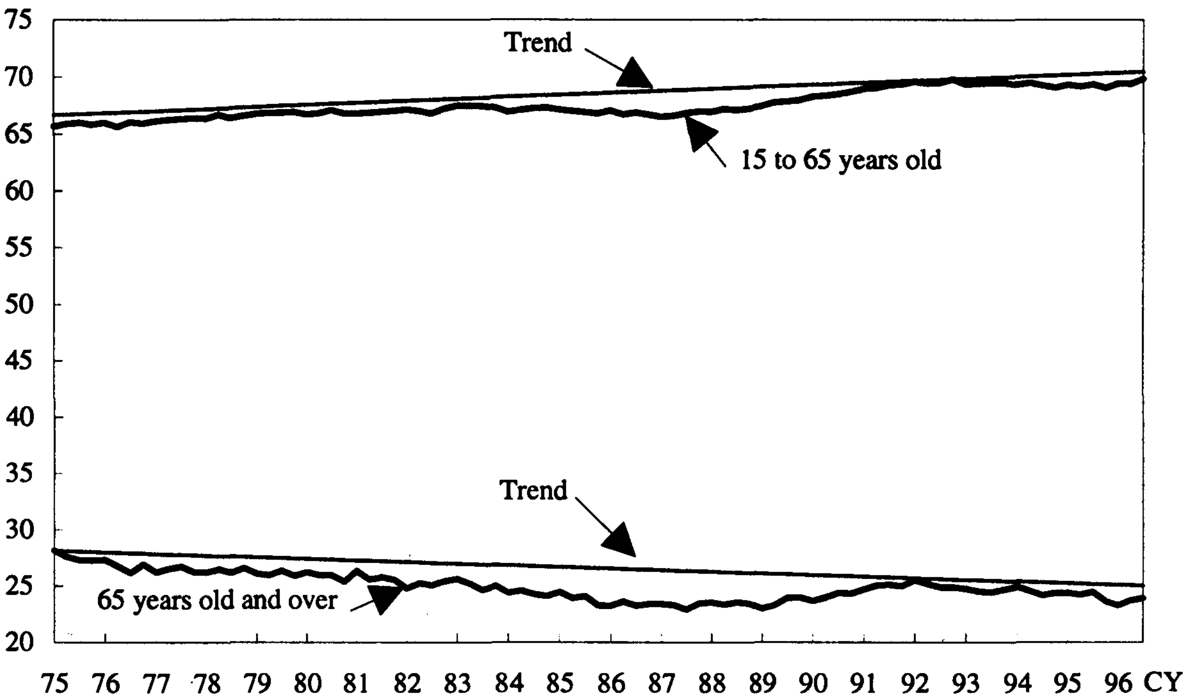
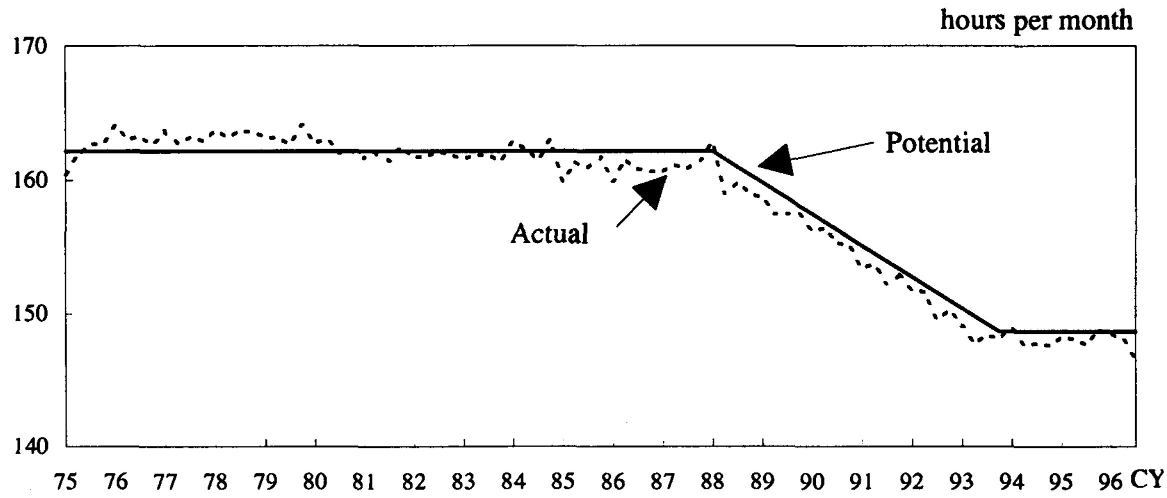


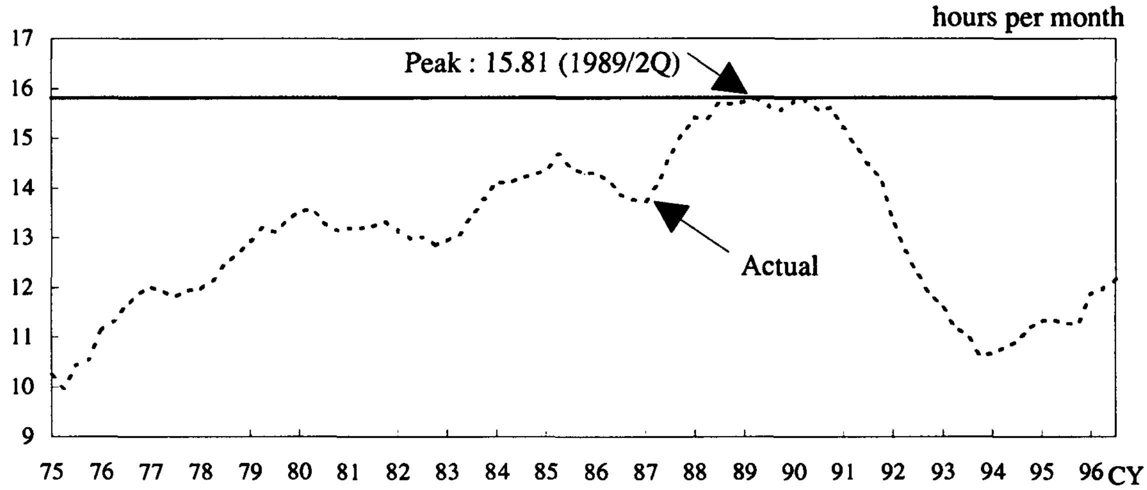
Figure 5

Working hours

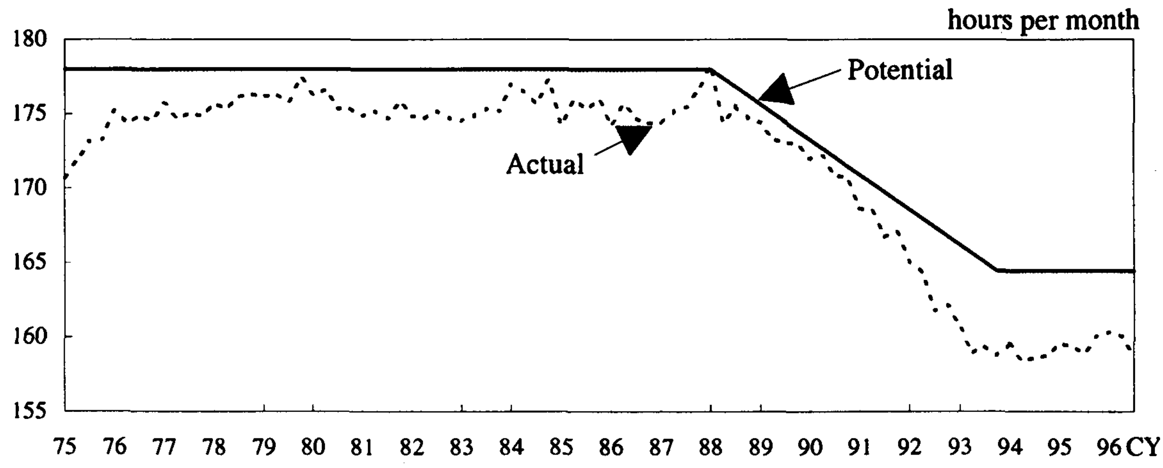
1. Scheduled working hours



2. Overtime working hours



3. Total working hours

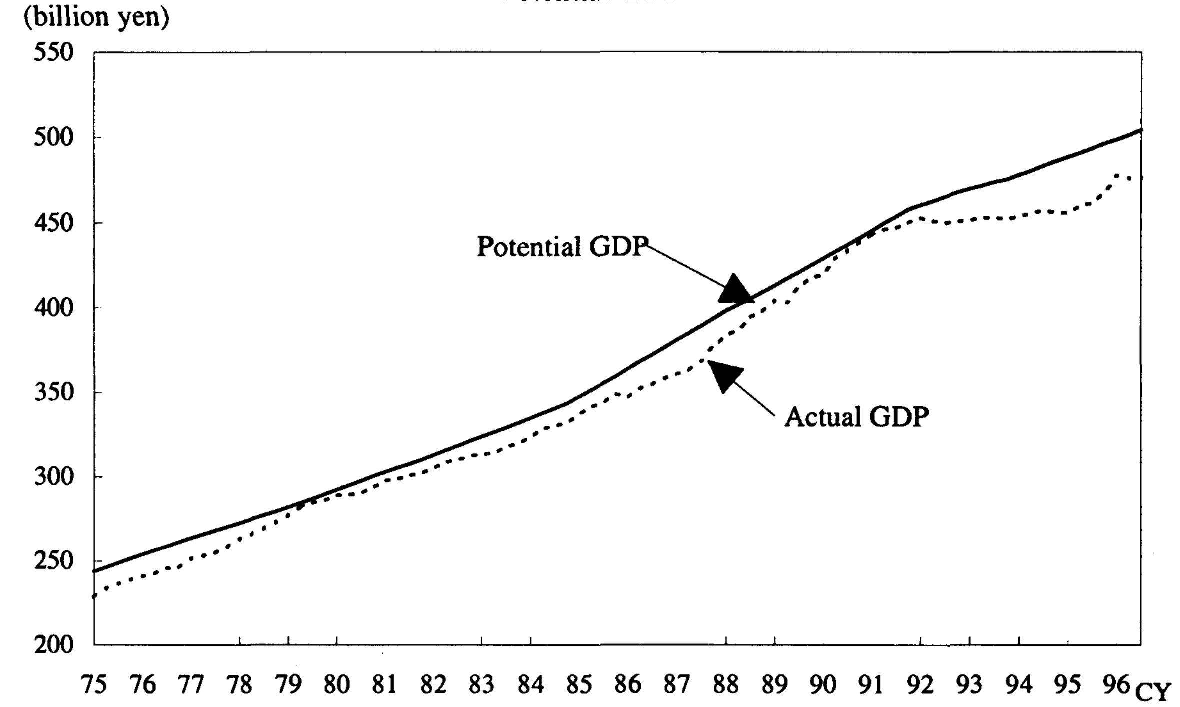


1.3 Potential GDP and GDP gap

We can calculate potential output by putting the potential amount of factor inputs into the production function. Potential GDP in Figure 6 is calculated in this way and the GDP gap in Figure 7 is defined as (actual GDP-potential GDP)/potential GDP.[2] As the figure shows, the GDP gap started to increase or deteriorate at the beginning of 1991 and continued to increase monotonically until the first quarter of 1995 when actual GDP deviated from the potential level by 7%. It tends to improve since the second quarter of 1995, although its level is still as high as 5%.

Figure 6

Potential GDP



To see what caused the deterioration of the GDP gap in the 1990s, we decompose changes in the GDP gap from the previous period into the following two components using the identity:

$$\Delta GDP\ gap = final\ demand\ /actual\ GDP - \Delta potential\ GDP\ /potential\ GDP - \Delta import\ /actual\ GDP \quad (3)$$

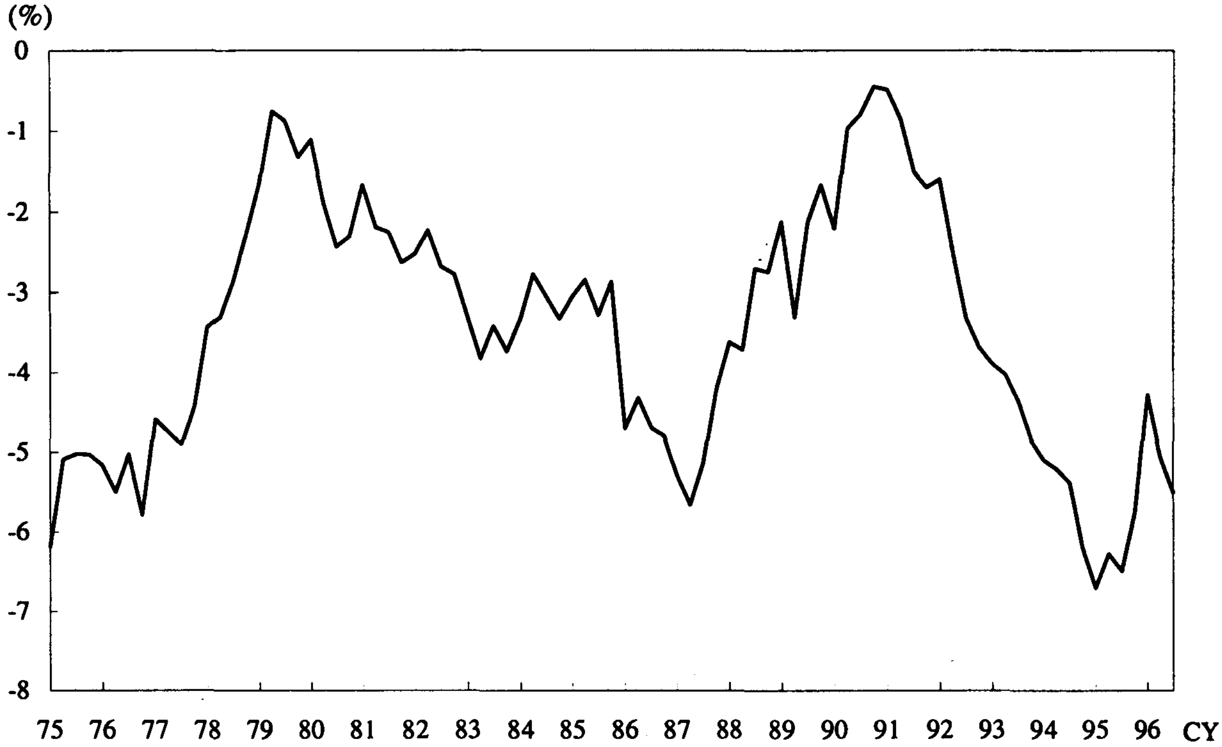
where the first and the second terms on the right-hand side are combined and called "final demand minus potential GDP factor" in Figure 8 and the third term "import factor".

As shown in Figure 8, the deterioration in the GDP gap in 1991 and 1992 was mainly explained by the first factor. That is, final demand, particularly domestic private demand, was weak as compared with potential output during this period, so that the GDP gap deteriorated. Fiscal and monetary stimulus starting in 1992 reduced the downward pressure from the import factor in and after

2 Note that the GDP gap obtained here and used in the rest of this paper is not the deviation of the gap from the natural rate. According to the Lucas supply function, the natural rate is defined as the level of the GDP gap where the actual rate of inflation coincides with the expected rate of inflation, as long as the supply shocks are negligible. Using the inflation forecasts made by private think-tanks each year as a proxy for the expected rate of inflation, we have found that the natural rate is about 3%.

Figure 7

GDP gap



1993. Instead, the downward pressure from the second factor has started to increase. This is due to an increase in imports from east Asian countries reflecting the appreciation of the yen as well as industrialisation in those countries. An important thing to note is that increase in imports cannot be explained by the movement of final demand. If all increases in imports had been induced by increases in final demand, they would have been neutral to the GDP gap, or the supply-demand condition in goods and service markets. In fact, final demand was too weak to induce any imports in this period, although it was gradually recovering. In this sense, the increase in imports in and after 1993 was almost independent of final demand, so that it increased the GDP gap significantly.[3]

Table 1

Selected examples of estimation of potential GDP

	Growth rate of potential GDP, in percentages				
	1991	**1992**	**1993**	**1994**	**1995**
Present paper	3.9	2.7	1.9	2.0	2.2
Economic Planning Agency	4.9	3.0	1.3	1.1	n.a.
Industrial Bank of Japan	4.1	4.1	3.1	n.a.	n.a.
Japan Development Bank	4.0	2.7	1.9	2.3	2.2
Bank of Tokyo-Mitsubishi	4.0	4.0	3.5	3.5	3.5
IMF	4.5	3.9	3.3	2.8	2.6
OECD	2.9	2.3	2.2	3.2	3.0

3 Table 1 shows the selected estimates of potential GDP and the GDP gap by various institutions, including the Economic Planning Agency and private think-tanks.

Figure 8

Decomposition of movements in the GDP gap

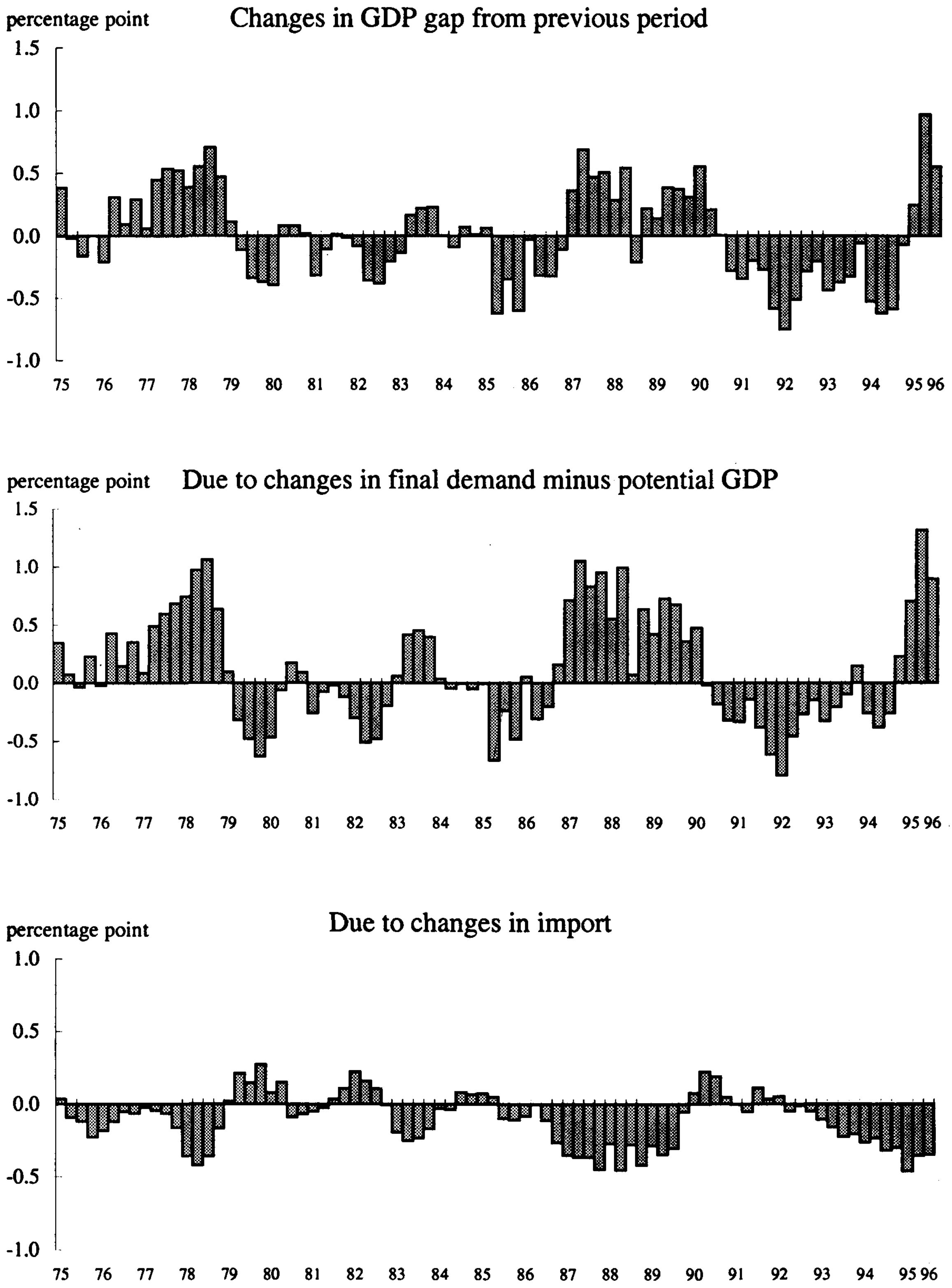
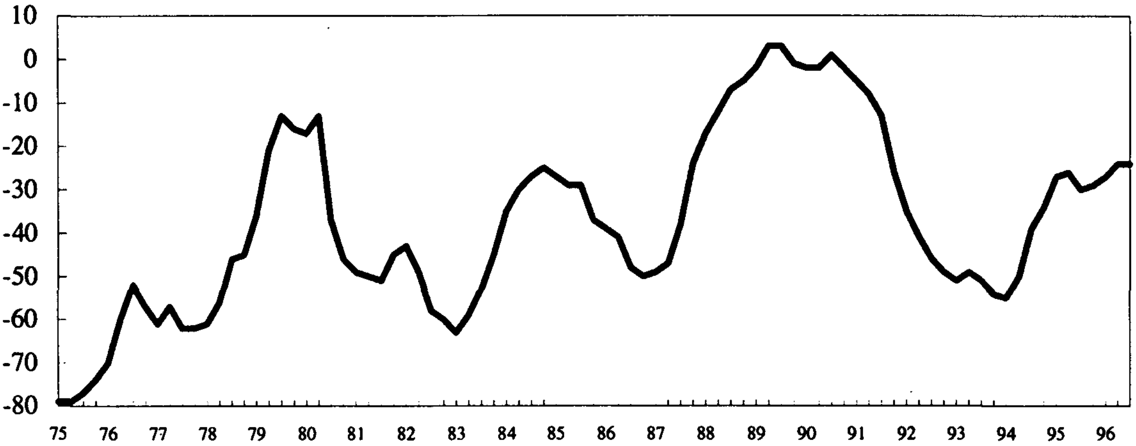


Figure 9

Indicators of the output gap

1. Diffusion index for "Supply/demand conditions for products"



Note:
Short-term Economic Survey of Principal Enterprises. Figures represent the percentage of "Excess of Demand" in total minus the percentage of "Excess of Supply".

2. Capital utilization rate

3. Unemployment rate

1.4 Other indicators of the output gap

Figure 9 (on the previous page) shows recent movements in various measures of slack in the economy: (i) the diffusion index, DI, concerning "supply/demand conditions for products" in *Short-term Economic Survey of Principal Enterprises*; (ii) the capital utilisation rate; and (iii) the unemployment rate. As clearly seen, all three measures commonly indicate the deterioration in the output gap during the period of 1990 to 1993. This is consistent with the movement of the GDP gap estimated above. However, the movements of those measures differ significantly in and after the beginning of 1994: the DI and capital utilisation rate have started to improve gradually since the beginning of 1994, while the unemployment rate and the GDP gap continued to deteriorate.[4]

We can point to the following reasons in explaining the different movements since the beginning of 1994. First, the difference in coverage: the DI and the capital utilisation rate represent slack in the manufacturing sector, while the GDP gap and the unemployment rate cover the whole economy. Second, the DI and the capital utilisation rate might fail to capture the impact of the rapid increase in imports from east Asian countries, which was the main source of deterioration of the GDP gap in 1994 and 1995. For example, correspondents of the *Short-term Economic Survey*, major enterprises from capital of over 1 billion yen, might not face severe competition from east Asian countries. It might be smaller enterprises whose products are more labour intensive that are exposed to competition from east Asia. The third possible reason is measurement error. In particular, we cannot deny the possibility of overestimating the potential growth rate, given that the growth rate of TFP we use in calculating the potential GDP is higher than the actual growth rate of TFP during this period.

2. Relationship between the output gap and inflation

2.1 Three hypotheses

There are three alternative hypothesis about the relationship between the output gap and inflation. First, the *level* of the output gap is related to the rate of inflation. The theory of general equilibrium tells us that the rate of change in each commodity price depends on the excess supply or demand for that commodity. By analogy, at the macro level this implies that the rate of change in the general price level, or the inflation rate, depends on the level of slack in the economy, or the output gap. A famous example of this type of relationship is the so-called Phillips curve, a stable relationship between the rate of change in wages and the unemployment rate.

Second, the level of the output gap might be related with the rate at which inflation rate increases, or the *acceleration* of the general price level. A famous example of this sort is the Non-Accelerating Inflation Rate of Unemployment (NAIRU). The rate of inflation continues to increase if the unemployment rate is below the natural rate, and continues to decrease if the unemployment rate is above the natural rate. It is only when the unemployment rate coincides with the natural rate that the rate of inflation remains unchanged. The NAIRU is a special case of the hypothesis of the natural rate of unemployment where the expected rate of inflation is equal to the rate of inflation in the previous period.

Third, the *change* in output gap might be related with the rate of inflation. For example, the general price level started to rise in the US economy just after the Great Depression reflecting the rapid shrinking of the output gap, with the level, measured by the GNP gap still being over 40%.

---

4 As for the unemployment rate, since a structural change is going on in the labour market, it might be inappropriate to consider such rise in the unemployment rate as entirely cyclical.

Some argue that this sort of relationship was observed in instances other than Great Depression and call it the "speed limit effect" (for example, see Romer (1996)).

In what follows, we will see which of the three hypothesis holds for the Japanese economy. Before proceeding further, however, let us briefly think about implications of each hypothesis for the future course of inflation in the Japanese economy. As we saw in Section 1, the current level of the GDP gap in Japan is as high as 5% and the current rate of inflation is near zero. Suppose someone predicts that the GDP gap will improve in 1997 as the economy recovers but the speed of improvement will be limited reflecting the weakness of final demand. What does this prediction imply for the future rate of inflation?

According to the first hypothesis, it is the level of output gap that is related to the rate of inflation. Therefore, those who believe in the first hypothesis predict that the rate of inflation will gradually rise as the GDP gap improves. For those who believe in the second hypothesis, the critical thing is to know whether the current level of the GDP gap is below or above the natural level. Needless to say, it is impossible to prepare any answer about the natural rate without close examination of the data; nevertheless, most economists will probably agree that the natural rate, if it exists, will be lower than 5%. If this is the case, the rate of inflation will continue to go down in the near future. Finally, according to the third hypothesis, it is not the level of the GDP gap but the speed of improvement that determines the rate of inflation. No matter how large the current and future level of the GDP gap, inflation will surely increase as long as the GDP gap improves.

In the rest of this section, we will investigate which of the three hypotheses holds for the Japanese economy through "eyeball econometrics" as well as simple regressions.

2.2 The Phillips curve

Figure 10 plots the level of the GDP gap and the rate of inflation measured by the CPI. It is clearly seen that there exists a stable trade-off between the two variables since 1980. In other words, we can observe something like the Phillips curve in the 1980s and the first half of the 1990s. It is surprising that we observe such a stable relationship between the two in spite of various shocks hitting the Japanese economy during this period such as the rapid appreciation of the yen after the Plaza Agreement in 1985, the asset price inflation in the late 1980s, and the burst of the bubble in the 1990s.

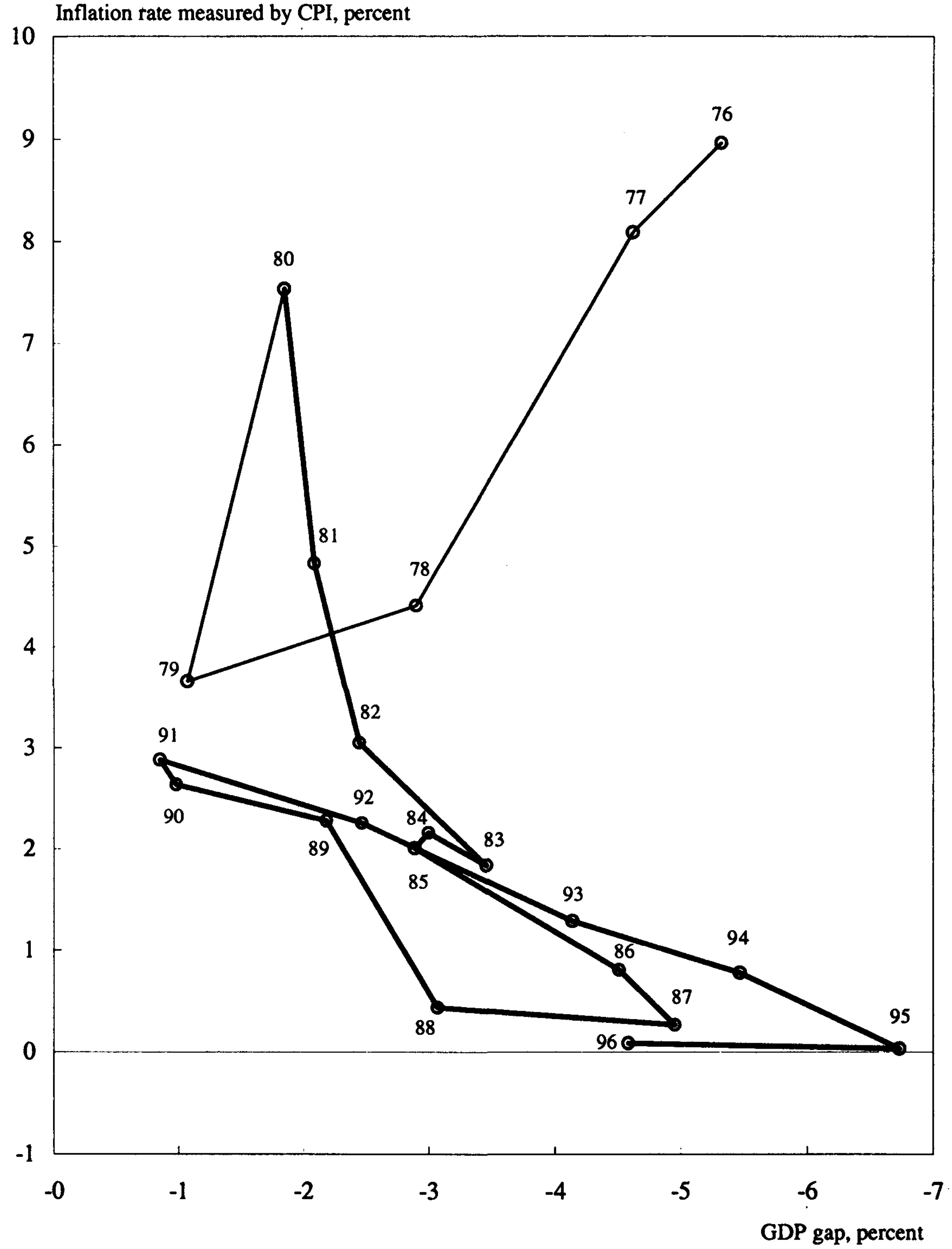
Those who believe in the natural rate hypothesis might argue that from 1980 to 1996 is too long a time range to observe a stable Phillips curve: the expected rate of inflation must have shifted at least several times during such a long period. We agree that there might not be a single Phillips curve but a set of multiple Phillips curves. For example, the curve in 1980-84 clearly differs from that in 1985-87 with respect to slope and intercept. Even if we take this point into consideration, however, it still seems surprising that the two variables are closely related with each other. An interpretation of this fact based on the natural rate hypothesis is that expectations about future inflation were relatively stable. In fact, various surveys, including the *Short-term Economic Survey* conducted by the Bank of Japan, commonly indicate that the expectations of firms and households have been relatively stable since the beginning of the 1980s.[5]

---

5 This is consistent with the observations that the money supply and nominal GDP have been moving together since the beginning of the 1970s, and that the money supply and the GDP deflator were also moving together up to the early 1980s, whereas a stable relationship has been missing since the mid-1980s. That is, according to the natural rate hypothesis, there is a trade-off between inflation and unemployment in the short run, during which the expected rate of inflation is kept constant; the Phillips curve is vertical in the long run when the expected rate of inflation coincides with the actual rate. In the 1970s, the inflation rate was so high that the Phillips curve was almost vertical. As a consequence, the rate of inflation was governed by the growth rate of money supply. In contrast, the expected rate of inflation was stable in the 1980s and 1990s, so that the rate of inflation was mainly determined by the GDP gap.

Figure 10

GDP gap versus CPI inflation

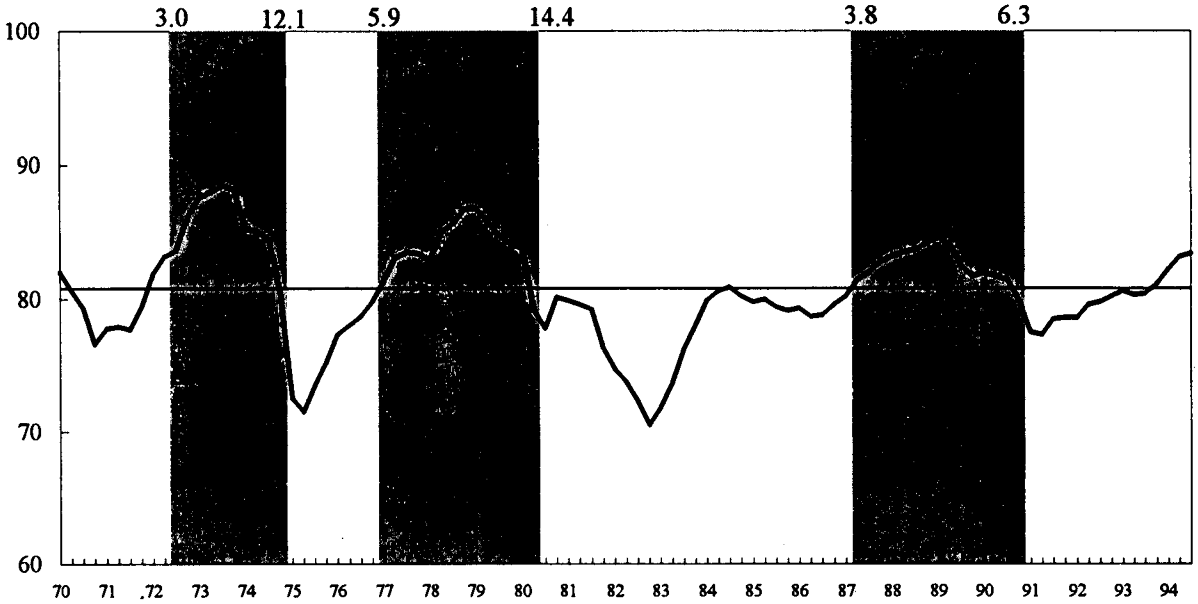


Notes: CPI is excluding fresh food and consumption tax. Figures are computed on calendar-year basis. The figure for 1996 is the percentage change from a year ago of the 1st semester 1996.

Figure 11

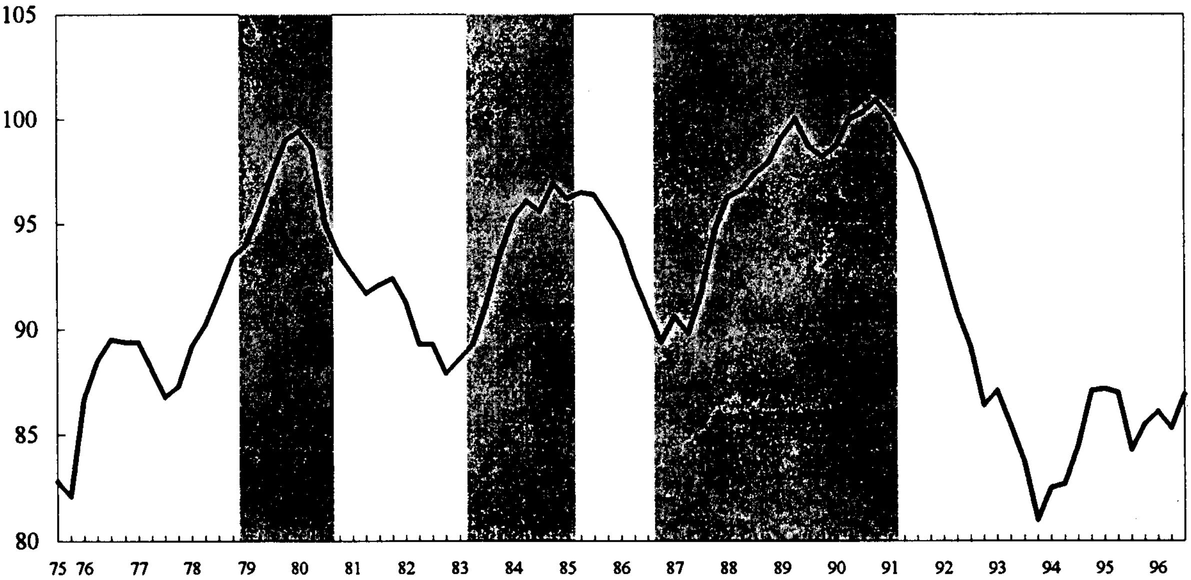
Capital utilisation rate versus acceleration of inflation

U.S.



Note : Shaded areas represent the periods during which the inflation rate increased.
Source : Garner,Alan,C. (1994). "Capacity Utilization and U.S. Inflation." *Economic Review*, Federal Reserve Bank of Kansas City. Vol.79,No.4,5-21.

JAPAN



Note: Shaded areas represent the periods during which the inflation rate measured by the TC component of the CPI excluding fresh food and consumption tax increased.

2.3 The NAIRU

Figure 11 shows the relationship between the acceleration of the CPI and the capital utilisation rate in the manufacturing sector. The shaded area represents the period during which the rate of inflation increased. If the relationship is of the NAIRU type, we should observe that the rate of inflation increases when the capital utilisation is above a specific level, or the natural level, and vice versa.

The upper panel, which is taken from Garner (1994), shows the relationship in the US economy. As clearly seen, there is a tendency that the CPI accelerates when the capital utilisation is above 80.2 and decelerates when it is below that level.[6] In contrast to this, as shown in the lower panel, we cannot observe such regularity for the Japanese economy: the NAIRU-type relationship between the output gap and inflation does not exist in Japan. Rather, looking more closely, it seems that the rate of inflation increases when the capital utilisation increases and vice versa. In other words, it is not the acceleration of CPI but the inflation rate of CPI that is related to the level of the capital utilisation rate. This is consistent with what we observed in Figure 10.

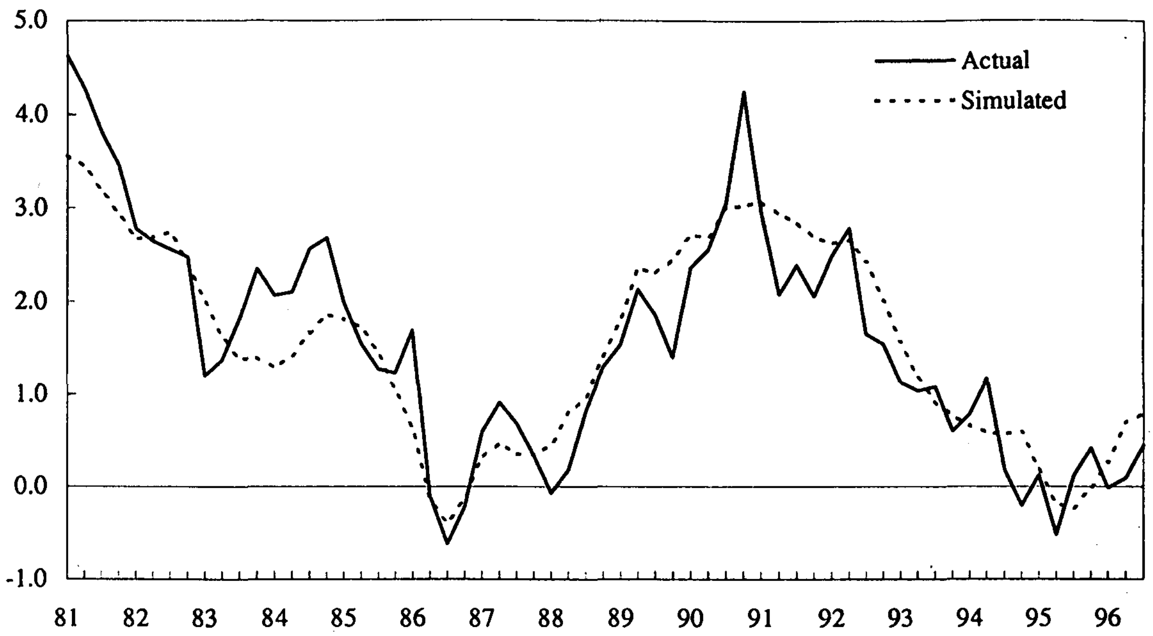
2.4 Estimation of the inflation equation

Examination of Figures 10 and 11 through "eyeball econometrics" seems to indicate that the rate of inflation is related to the level of the output gap in Japan. To formalise this finding, we regress the rate of inflation (π_t) on the lagged rate of inflation (π_{t-1}), the output gap (g_{t-1}), change in the output gap (Δg_{t-1}), and the rate of change in import prices (m_t). That is,

$$\pi_t = a_0 + a_1 \pi_{t-1} + a_2 g_{t-1} + a_3 \Delta g_{t-1} + a_4 m_t \quad (4)$$

Figure 12

Dynamic simulation of the inflation equation (CPI excluding fresh food and consumption tax versus the GDP gap) estimated over 1981Q1 to 1996Q3



6 Note that the rate of inflation has not significantly increased since 1994 although the capital utilisation rate is well above 80.8. In this sense, the relationship of the NAIRU type seems to disappear even in the US economy.

We will use changes in CPI total and CPI goods as indicators of π_t and the GDP gap, the capital utilisation rate and the unemployment rate for g_{t-1}. The sample period is 1981Q1 to 1996Q3 and the estimation method OLS.

If the first hypothesis holds, a_2 should be significantly different from zero and a_1 should be close to zero. If the second hypothesis holds, a_2 should be significantly different from zero and a_1 should be equal to unity. Finally, if the third hypothesis holds, a_3 should be significantly different from zero.

Table 2 shows the results of the regressions.[7] Numbers in parentheses represent standard errors. The results are summarised as follows: (i) a_2 is different from zero at the 5% significance level in all cases; (ii) a_1 is different from zero at the 5% significance level in all cases except the combination of CPI total and the unemployment rate; (iii) in all cases, the null hypothesis that a_3 is zero cannot be rejected at the 5% significance level. Altogether, we can conclude that the data support the first hypothesis but contradict the other two. This is consistent with the finding obtained through "eyeball econometrics".

Table 2

Estimation of inflation equation

Output gap measured by:		**Inflation measured by:**	
		CPI total	**CPI goods**
GDP gap	a_0	1.64 (0.33)	2.70 (0.61)
	a_1	0.54 (0.08)	0.25 (0.11)
	a_2	0.27 (0.06)	0.60 (0.15)
	a_3	-0.10 (0.11)	0.21 (0.30)
	a_4	0.02 (0.01)	0.04 (0.02)
	$\bar{R}^2$	0.83	0.51
	S.E.	0.50	1.44
	D.W.	0.76	-1.48
Unemployment rate	a_0	2.52 (0.87)	5.80 (1.87)
	a_1	0.66 (0.08)	0.36 (0.11)
	a_2	-0.78 (0.30)	-2.06 (0.71)
	a_3	0.99 (0.79)	2.49 (2.22)
	a_4	0.02 (0.01)	0.04 (0.02)
	$\bar{R}^2$	0.80	0.43
	S.E.	0.54	1.55
	D.W.	0.89	-0.98
Capital utilisation rate	a_0	-3.25 (1.37)	-9.10 (4.03)
	a_1	0.73 (0.06)	0.43 (0.11)
	a_2	0.04 (0.02)	0.10 (0.04)
	a_3	-0.05 (0.05)	-0.04 (0.15)
	a_4	0.02 (0.01)	0.04 (0.02)
	$\bar{R}^2$	0.80	0.41
	S.E.	0.55	1.59
	D.W.	0.07	-1.94

Notes: Numbers in parenthesis represent standard errors. The rate of inflation is measured by the annualised rate of change from the previous quarter of the TC component of CPI total or CPI goods. Equations are estimated by OLS. See descriptions in the text for details.

[7] Figure 12 shows a dynamic simulation of the estimated equation which clearly performs well.

3. The cost of disinflation

As a concluding remark, let me briefly discuss the cost of disinflation the Japanese economy has experienced since the burst of the bubble. According to the literature, including Krugman (1996) and Fischer (1996), disinflation from, say, 5 to 0% is more costly in the terms of the sacrifice ratio than disinflation from, say, 10 to 5%. This is because the nominal wage has some downward rigidity so that the real wages rise when the rate of inflation approaches zero, leading to a decrease in the demand for labour and a higher rate of unemployment. According to their argument, since the cost of disinflation is higher than the benefit of such disinflation when the rate of inflation is close to zero, central banks should not target zero inflation. They should target, say, 3% or moderate inflation.

Their theoretical reasoning as well as the policy implication seems crystal clear, but their argument lacks empirical evidence because inflation rates have been far from zero in almost all industrial countries, at least during the post-war period.[8] In this circumstance, the experience of Japan since the burst of the bubble seems quite valuable from the point of view of evaluating the cost of moving to zero inflation.

3.1 Is the real wage too high?

According to the argument for moderate inflation, zero inflation is costly because of rigid nominal wages. It is not easy to test whether nominal wages are rigid or not, and this is beyond the scope of this paper. But, if the argument is correct, we should observe a rise in real wages as the rate of inflation approaches zero. More precisely, we should observe that real wages become too high as compared with labour productivity.

Based on this understanding, we compare the real wage and labour productivity in Figure 13. As the figure shows, dots are on or near the 45-degree line in the 1980s which means that the real wage has tended to comove with labour productivity during this period.[9] In contrast, dots deviate significantly from the 45-degree line in the 1990s. In this sense, the real wage has been too high during this period. By closer inspection, a similar phenomenon is found in the early 1980s when the Japanese economy was also weak.[10] These two observations seem to suggest that, when the growth rate of output is low, firms try to reduce labour inputs by first cutting overtime working hours and then the number of employed persons. But such adjustments usually take time, particularly the second, thereby leaving labour inputs too high as compared with the level of output during the adjustment process. As a result, the growth of labour productivity goes down. In this situation, firms try to control real wages in accordance with the lower labour productivity, but this is also a difficult task and takes time to complete. As a consequence, dots deviate from the 45-degree line during a recession.

One thing special to the first half of the 1990s is that the rate of inflation has been very low. We cannot deny the possibility that the lower rate of inflation makes it even harder for firms to control real wages during the recession.

---

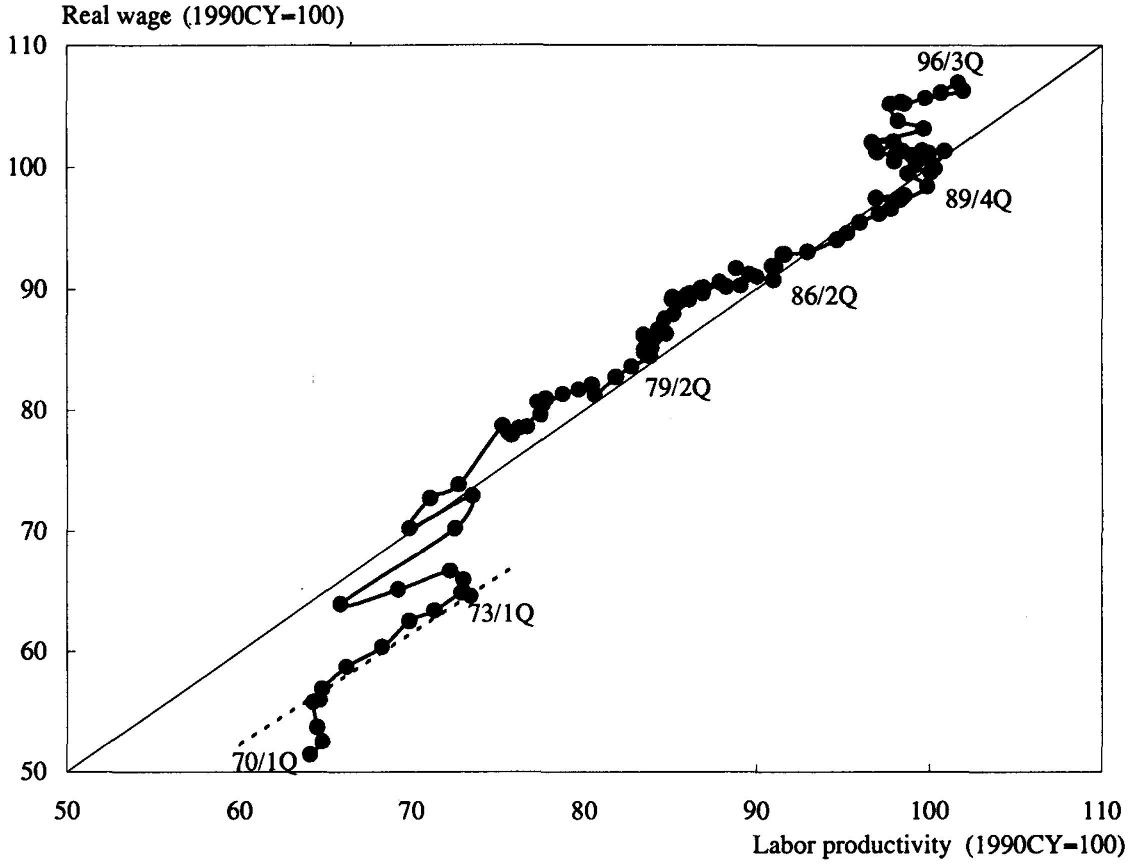
8 Needless to say, if we go back to the Great Depression or the period of the gold standard, you might find ample instances of zero or below-zero inflation.

9 Note that both the real and labour productivity are indices with 1990CY=100. Hence, the fact that dots are on the 45-degree line does not necessarily mean the level of real wages and the level of labour productivity coincide.

10 Another episode is the period 1973-74 when dots deviate from the 45-degree line of the 1970s. In this period, the rate of inflation of over 20% made expectations of inflation very unstable. People started to expect higher and higher inflation and nominal wages rose at a very high speed. In fact, the growth rate of wage outweighed the actual inflation rate. As a result, real wages rose while the growth of labour productivity was very low. Given this story, the deviation from the 45-degree line in 1973 and 1974 can be interpreted as caused by an overshooting of expectations.

Figure 13

Real wage and labour productivity



Notes: Real wage = nominal wage/GDP deflator. Labour productivity = real GDP * nominal wage/compensation of employees.

3.2 Is the Phillips curve flatter?

The observation that the real wage is too high is consistent with the argument for moderate as opposed to zero inflation. The next thing to check is whether the high real wage led to a rise in the unemployment rate. As we saw in Figure 9, the rate of unemployment has indeed been high since the burst of the bubble. But, as we noted in footnote 4, this is partly due to a structural change in the labour market; therefore, it might not be appropriate to attribute the rise in unemployment to higher real wages alone.

A more straightforward way to see the consequence of higher real wages is to check the slope of the Phillips curve. If a rise in the real wage reduces the demand for labour significantly, the slackness of the economy, including unemployment, must increase. If this is the case, the Phillips curve we observed in Figure 10 should be flatter in the first half of the 1990s than before. Looking at Figure 10 from such a point of view, we can say that the slope of the curve is clearly flatter in the period 1991-95 than in 1980-83. On the other hand, no significant change is observed between 1985-87 and 1991-95. What do these two observations mean? First, since the rate of inflation in both 1985-87 and 1991-95 was between zero and 3%, there is no reason why the slope should differ between these periods. Therefore, the second observation is exactly what the theory predicts. The first observation is more interesting. If we take this fact as it is, the Phillips curve is flatter in the 1990s and disinflation from 3% (in 1991) to zero (in 1995) is, therefore, more costly than disinflation from 7.5% (in 1980) to 1.8% (in 1983).

Figure 14

CPI inflation and nominal interest rate

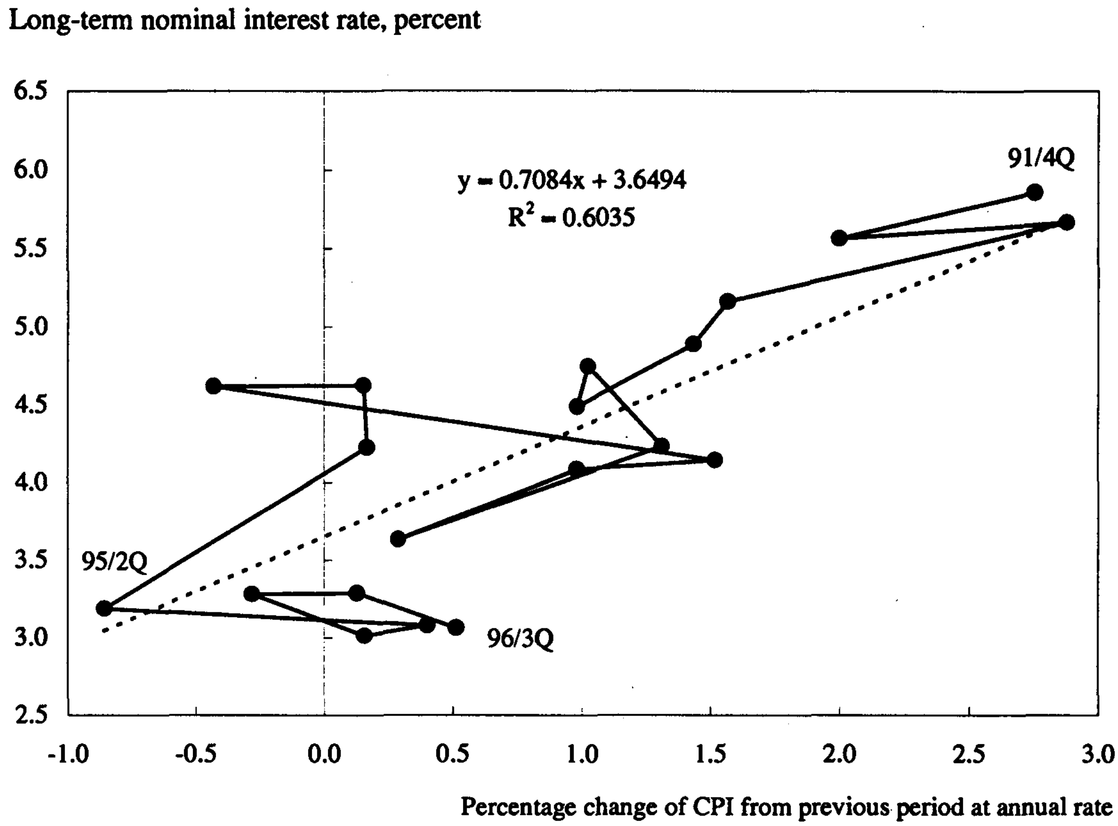
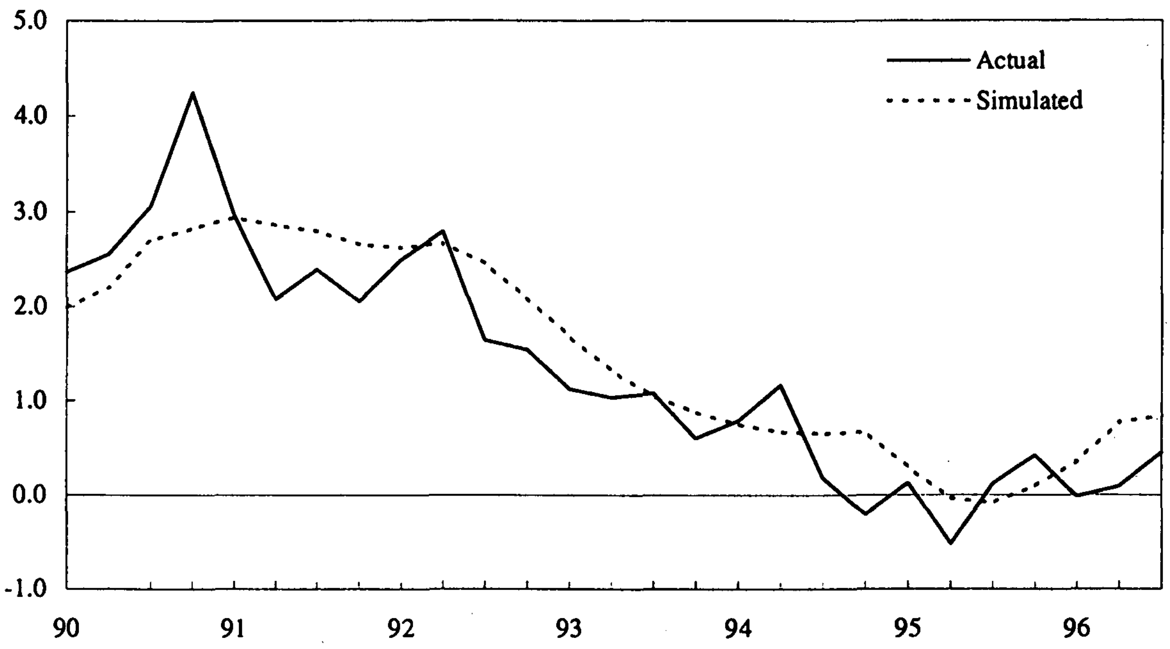


Figure 15

Dynamic simulation of the inflation equation (CPI excluding fresh food and consumption tax versus the GDP gap) estimated over 1981Q1 to 1989Q4



One caveat is that the CPI inflation rates in 1980 and 1981 were governed by oil prices. Obviously, it is not appropriate to look at the slope of the Phillips curve when a cost-push factor like higher oil prices affects the price index. In such a case, it is important to control for the effects of supply-side factors. Based on this understanding, we have conducted a simple experiment using the inflation equation estimated in Section 2.4. That is, we have reestimated the inflation equation using data for the 1980s and conducted out-of-sample forecast. If zero-inflation is costly, we should observe *underprediction*: that is, the forecast values should be lower than actual values. However, as shown in Figure 15, the forecast values are consistently higher than actual. In other words, when we properly control for the supply-side factors, the relationship between the rate of inflation and the GDP gap during the post-bubble period remains the same as that of the 1980s. In this sense, we cannot find any evidence supporting the hypothesis that the sacrifice ratio goes up as the rate of inflation approaches zero.

3.3 Future works

In sum, the evidence is mixed: the data seem to suggest that the disinflation since the burst of the bubble in Japan has caused higher real wages; at the same time, however, we cannot find strong evidence of a higher sacrifice ratio during that period.

This is just the beginning of research on this topic and lots of things need to be done to evaluate the cost of zero inflation.[11] For example, we need to investigate further why real wages remained high during the post-bubble period and whether this is an immediate and direct consequence of zero inflation. Also, further analysis of the downward rigidity of nominal wages will be needed.[12]

References

Fischer, Stanley (1996): "Why are Central Banks Pursuing Long-run Price Stability?" Paper presented at the Federal Reserve Bank of Kansas City Symposium, *Achieving Price Stability*.

Garner, Alan C. (1994): "Capacity Utilization and US Inflation". *Economic Review*, Federal Reserve Bank of Kansas City, Vol.79, No.4, pp. 5-21.

Kimura, Takeshi and Kazuo Ueda (1997): "Downward Nominal Wage Rigidity in Japan: Is Price Stability Costly?" *Working Paper*, Bank of Japan, May.

Krugman, Paul (1996): "Stable Prices and Fast Growth: Just Say No". *The Economist*, 31st August.

Romer, Christina D. (1996): "Inflation and the Growth Rate of Output". *NBER Working Paper* No. 5575.

---

[11] The advocates of moderate inflation are also concerned that zero inflation tends to lead to a higher real interest rate. For example, suppose a central bank, struggling with recession, wants to create a negative real interest rate. If the rate of inflation is zero, it is impossible to create a negative real interest rate because nominal interest rates cannot be negative. Also, if deflation is going on, the real interest rate remains high or continues to rise even if the nominal interest rate is very close to zero. From this point of view, the relationship between the long-term nominal interest rate and the rate of inflation is shown in Figure 14. As immediately seen, there is no evidence that the long-term nominal interest rate remained too high in comparison with the rate of inflation.

[12] See Kimura and Ueda (1997) as an example of this sort of study for Japan.

Comments on: "Output gap and inflation: the case of Japan" by Tsutomu Watanabe

by Steve Kamin

Mr. Watanabe's contribution is a clear, straightforward, and interesting analysis of the behaviour of inflation in Japan. During the 1980s, Japanese inflation was well below the OECD average, and more recently, Japan experienced the relatively novel phenomenon of price *deflation*. Under these circumstances, it is interesting to see whether Japanese inflation might be consistent with a standard Phillips-Curve model such as Mr. Watanabe estimates for Japan.

To review, the paper first estimates a production function for Japan and then, based on hypothetical values for inputs, calculates a measure of potential GDP. Using the resultant measure of the output gap for Japan, Mr. Watanabe estimates a standard Phillips-Curve type equation, augmented with a measure of import prices. Finally, he discusses the applicability to Japan of the hypothesis that as inflation declines, money illusion may make it increasingly difficult to achieve zero inflation itself.

I found the modelling strategy employed by Mr. Watanabe to be reasonable, and his basic results were plausible as well. He finds the level of the output gap to affect inflation significantly, as one might expect, but identifies no "speed limit" effects – that is, the change in the output gap does not affect inflation. In contrast to findings for some other countries, including the United States, he finds the coefficient on the lagged inflation rate in the Phillips Curve equation to be significantly less than one, suggesting that the long-run Phillips Curve for Japan is less than vertical. This should not dismay devotees of rational expectations theory: rational expectations behaviour tends to emerge when inflation is high and variable, and given the relatively low inflation enjoyed by Japan in recent years, it is not surprising that past inflation is not fully and immediately incorporated into current inflation.

Mr. Watanabe's approach towards calculating potential output – estimating an actual production function – may be superior to "black box" methods such as fitting a curve or filter to actual GDP. However, it should be cautioned that because certain crucial questions about the Japanese economy remain unresolved, Mr. Watanabe is forced to make modelling decisions that are only slightly less "ad hoc" than, say, fitting an HP filter to the data. For example, it remains a mystery why Japanese output grew so rapidly during the "bubble economy" period of the late 1980s, and then has remained so stagnant thereafter. Observers have argued that the slump of the 1990s reflects a ratcheting down of Japan's potential growth rate, but no convincing rationale for the timing of this decline has been offered. In order to capture this phenomenon, Mr. Watanabe allows for a greater rate of total factor productivity growth during the late 1980s than either before or after. Since the bubble economy period probably was demand rather than supply-driven, however, the justification for this procedure is unclear; it probably was required in order to keep the measured output gap from becoming too large and positive during the bubble economy period, and then becoming too large and negative during the 1990s.

A second concern raised by the paper is the role of import prices in the recent slowdown in Japanese prices. While import prices (measured in yen) are estimated to have a statistically significant impact on Japanese inflation, they are not discussed sufficiently in the text of the paper. Considering the tremendous appreciation of the yen during the 1990s, the fall in yen import prices would seem to be a good candidate for explaining much of the decline in inflation. It would have been useful to see a counterfactual simulation of the path of inflation during the 1990s, holding import prices constant.

Finally, Mr. Watanabe's discussion of the costs associated with zero inflation is apt, considering that Japan recently has experienced negative inflation. However, more could have been done with the data to test the hypothesis that, as price growth approaches zero, it becomes more difficult to reduce nominal wage growth accordingly. The paper shows that real wages, relative to productivity, have risen as of late, but it is difficult to tell whether this reflects lower inflation or other, unrelated factors such as lower productivity growth due to labour hoarding. An attempt also is made to determine whether the Phillips Curve has become flatter in recent years, but it is tough to make sense of the spaghetti squiggle of lines in Figure 10. Finally, the out-of-sample forecast presented in Figure 15, showing that the inflation equation estimated during the 1980s continues to track relatively well in the 1990s, *does* present prima facie evidence against the "costly disinflation" hypothesis. However, it is possible that other factors not included in the equation, such as the spread of discount stores, may have acted to offset inflationary pressures resulting from nominal wage rigidities.

It might have been more informative to have performed a direct test of the hypothesis that wages or prices behave differently at different levels of inflation. For example, one could examine the linkage between wage growth and inflation at different levels of inflation:

$$\Delta[\mathit{Wage\ Growth}]_t = \alpha + \beta\Delta[\mathit{Inflation}]_{t-1} + \eta\Delta[\mathit{Dummy}]*[\mathit{Inflation}]_{t-1} + \delta\Delta[\mathit{Output\ Gap}]_{t-1}$$

Dummy: 1 if inflation < 3%, else 0

Δ: difference operator

This equation can be used to test the hypothesis that when inflation is sufficiently high – say, over 3% – changes in inflation, for a given output gap, translate one-for-one into changes in wage growth; this would be reflected in an estimated β of unity. On the other hand, if, as inflation fell below 3%, wage growth declined by a lesser amount as a result of money illusion, the estimated η should be negative.

One could perform an analogous regression in Phillips Curve space to determine whether changes in the output gap had a smaller impact on inflation as inflation declined:

$$\Delta[\mathit{Inflation}]_t = \alpha + \beta\Delta[\mathit{Inflation}]_{t-1} + \delta\Delta[\mathit{Output\ Gap}]_{t-1} + \eta\Delta[\mathit{Dummy}]*[\mathit{Output\ Gap}]_{t-1}$$

With this equation, if progressively larger increments to the (negative) output gap are required to reduce inflation as it nears zero, the estimated η should be negative, as above.

A systems approach to the determination of the NAIRU, inflation and potential output in Austria

Friedrich Fritzer and Heinz Glück[1]

Introduction

After decades of very low unemployment rates, tensions in the labour market also appeared in Austria by about the middle of the 1980s. Under the influence of the opening of the Eastern European countries, Austria's accession to the EU, globalisation and other shocks, the discussion has intensified whether and to what extent these developments may have contributed to the ongoing rise of unemployment and whether they will continue to do so. Problems of this kind are frequently discussed in terms of the natural rate of unemployment or the NAIRU (non accelerating-inflation rate of unemployment), which relate labour market concepts to inflation and potential output. As is well known, the NAIRU seems have risen in the last few years – not only in Austria, but in many other countries. From a policy perspective, it is of foremost interest whether the NAIRU can be sustainably reduced through deliberate economic policy measures by working, for instance, directly through labour market institutions and the wage bargaining process (OECD 1994). However, Ball (1996) pointed out that the disinflation policy of the last years has caused a considerable outward-shift of the NAIRU in many countries (i.e. the stability of the price level is now associated with higher unemployment than before).

During the 1970s, Austria's unemployment rate fluctuated within narrow margins around 2%. In 1980, unemployment started to accelerate, reaching 6.6% in 1995. Immediate questions are: what were the reasons for this development and what can we do about it?

In case of Keynesian unemployment, we should call for measures to stimulate effective demand; i.e. higher public spending or a looser monetary policy. At the moment, these measures are not very popular because of the binding constraints of the Maastricht criteria; and if we scan the data we cannot actually find any indication that the rise in unemployment is linked to too restrictive a course in terms of public spending or monetary policy. Budget deficits have been rising since the 1980s, and the downward trend in inflation already started in the mid-1970s when unemployment was pretty stable at about 2%.

Classical unemployment would instead call for measures to reduce the rigidity of real wages. In the analysis below, we will not deal with this case because the wage bargaining process among the social partners in Austria is meant to prevent unemployment.

Other sources of unemployment may be changes in production patterns which affect the demand and supply side of the labour market. Demographic developments, generous unemployment benefits, non-wage labour costs are some of the factors which drive the equilibrium unemployment rate. Unfavourable shifts in these determinants will increase the natural rate of unemployment and induce a parallel movement of actual unemployment or an ever accelerating inflation rate. If this is the case, the only wise response to high unemployment lies in the removal of structural impediments to the way the labour market works.

This paper tackles the questions raised above. After a short review of the results of some recent studies on Austria in Section 1, we specify and estimate single equations for wages, prices,

[1] The authors are economists at the Oesterreichische Nationalbank, Economic Analysis Division. The views expressed in this paper are those of the authors and not necessarily of the institution with which they are affiliated.

output and unemployment in Section 2. In Section 3, this system of equations is estimated by 3SLS, and in the final section we try to draw some conclusions from the results.

1. The discussion in Austria

This section gives a short – and certainly incomplete – review of the discussion on wage-price relations, the Phillips curve, potential output and the NAIRU in Austria.

As far as the more recent investigations of these issues in Austria are concerned, we start with Wörgötter's (1983) contribution, in which he tries "... to give an explanation for the astonishing stability of the relationship between wage and price increases on the one side and unemployment rates on the other side in Austria from the end of the post-war recovery in 1955... Contrary to the experience in other industrialised countries no outward shift of the Phillips curve could be observed during the 1970's...". Wörgötter develops an equation system for the simultaneous determination of wages and prices and concludes from the results that the "... influence of social partners on the determination of wages and prices in Austria allows for a stable long-run empirical relationship between inflation and unemployment... The self-adjusting behaviour of social partners in Austria...removes this constraint on economic policy which can rely on the endogenous stabilisation of wages and prices, without being forced to apply an anti-inflationary fiscal and monetary policy."

In the context of monetary and exchange rate policy, Handler (1989) took a somewhat different view. He feels that, due to external influences (price shocks), the Phillips curve for Austria cannot be interpreted as a trade-off between inflation and unemployment. Therefore, monetary policy decisions should be only loosely based on this relation. Moreover, he did not find any empirical proof that the permanent dampening of the inflation rate through the hard-currency option had negative effects on the real economy and on unemployment. Thus the Phillips curve is no constraint for a stability-oriented monetary policy.

Pichelmann (1989), however, showed that the NAIRU doubled during the fifteen years preceding his investigation. He interprets this rise as evidence of "... an increase of distributional pressures in the Austrian economy" and addresses the question whether the Austrian political and organisational system may have lost some of its flexibility. Employing a simple macro-economic model of wage-price formation, he establishes that both the increase in employers' contributions to social security and the rising share of long-term unemployment associated with higher overall unemployment have exerted pressure on real product wages and thus exerted upward pressure on "the equilibrium rate of unemployment". Though hysteresis will have played an important role, the rise of NAIRU was regarded as evident.

Wehinger (1990) pointed out the problem that potential output and the natural rate of unemployment are frequently determined independently or only in a recursive way, and successfully tries a simultaneous approach. He identifies some marked output gaps between 1967 and 1988, but finds that they are slightly smoother compared to calculations using simpler methods.

Gartner's study (1995) represents a first attempt by Oesterreichische Nationalbank to introduce the concepts of potential output and the output gap into monetary analysis. Trend output was determined by means of the Hodrick-Prescott filter, and Granger causality tests were used to check whether the values for the output gap could contribute to forecasting inflation, which was confirmed.

Hahn and Rünstler (1996) try to combine astructural (e.g. the Hodrick-Prescott filter) and structural approaches in determining potential output, the latter being deduced from wage-price equations. They find that actual GDP in Austria was almost always below potential output since the beginning of the 1990s, and that the average length of cycles was 28 quarters, with the cyclical component of unemployment lagging behind the GDP cycle by about three quarters.

Pichelmann and Schuh (1996) point out some problems inherent in the concept of NAIRU. They survey a number of hysteresis-mechanisms which could lead to permanent shifts of equilibrium unemployment over time, implying that a long-run NAIRU may not even exist. Other problems are that empirically derived NAIRU estimates are highly dependent on model specifications, and that a considerable amount of statistical imprecision is inherent in the estimates. Therefore, they argue, any policy conclusions and recommendations drawn from the NAIRU must be judged with utmost care.

2. Single equation results

With these caveats in mind, we approach the task of finding some empirical evidence for the items involved. As mentioned, it is necessary to simultaneously estimate an equation system like the one below[2] in order to get consistent parameter estimates for the determination of wages, prices, potential output, and unemployment:

$$w = \alpha_0 + \alpha_1 p^c + \alpha_2 q^{pot} + \varepsilon_\omega \quad (1)$$

$$\Delta w = \beta_1 \Delta p^c + \beta_2 \Delta p^c_{-1} + \beta_3 \Delta\left(U - U^{nat}\right) + \beta_4 \Delta\Delta\left(U - U^{nat}\right) + \beta_5 \Delta\left(p^c - p\right) + \beta_6 \Delta\left(p^c - p\right)_{-1} + \beta_7 \varepsilon_{\omega-1} + \varepsilon_{\Delta\omega} \quad (2)$$

$$\Delta U = \gamma_1 \Delta\left(y - y^{pot}\right) + \gamma_2 \Delta\left(y - y^{pot}\right)_{-1} + \gamma_3 \Delta\left(\Delta app + \Delta app_{-1}\right)/2 + \gamma_4 \Delta nwlc + \gamma_5 \Delta union + \varepsilon_{\Delta u} \quad (3)$$

$$\Delta p = \delta_0 + \delta_1\left(\Delta ulc + \Delta ulc_{-1} + \Delta ulc_{-2}\right)/3 + \delta_2 \Delta p^m + \varepsilon_{\Delta p} \quad (4)$$

$$y - 0.75h - 0.25k = \zeta_0 + \zeta_1 T + \zeta_2\left(y - y^{pot}\right) + \varepsilon_y \quad (5)$$

Equation (1) reflects the target-real-wage-bargaining model. In such an environment, unemployment in excess of the natural rate can coexist with stable wage and price inflation. Equation (2) is an equation for the short-run dynamics of wage inflation. The growth of wages Δw relates to consumer price inflation Δp^c, the change in the unemployment gap $\Delta(U-U^{nat})$, the growth of the gap between consumption and output prices $\Delta(p^c-p)$ and the lagged residuals from equation (1), which can be interpreted as the deviations of real wages from their target level. The change of the unemployment rate (equation (3)) is determined by the deviation of real output from its potential value y^{pot},[3] the growth of non-wage labour costs $\Delta nwlc$, the growth of union density $\Delta union$ and the growth of the share of apprentices as a percentage of the labour force Δapp. The evolution of prices (equation (4)) is estimated from unit labour costs Δulc (the divergence between the growth rates of wages and potential productivity) and from the growth rate of import prices Δp^m. The production function (equation (5)) is solved for multi-factor productivity (y–0.75h–0.25k) which is determined by a time trend T and the output gap.

The preliminary specification search has to be performed by single-equation estimates using OLS. This is done in the following paragraphs.

---

2 This approach goes back to Coe and Krueger (1990), who apply the target real wage bargaining model to the German labour market.

3 In the single equation estimation all potential values have been proxied by a Hodrick-Prescott filter.

2.1 The wage equation

Starting with the general wage formulation of equation (2), we tried to explore some additional and specific issues, which did, however, they did not contribute very much to the overall picture. We tried to include variables or proxies for unionisation, formulated more complex hypotheses for price expectations, and also experimented with error correction terms. Finally, we ended up with an estimation of the following form:

$$\Delta w = 0.882\Delta p^c + 0.819\Delta p^c_{-1} - 3.170\Delta\left(U - U^{nat}\right) + 1.916\Delta\Delta\left(U - U^{nat}\right) - 0.941\Delta\left(p^c - p\right)$$

(5.39) (4.35) (-4.05) (3.59) (-3.88)

$$-0.747\Delta\left(p^c - p\right)_{-1} - 0.155\varepsilon_{\omega-1} + \varepsilon_{\Delta\omega} \qquad (2')$$

(-3.02) (-2.24)

Numbers in parentheses are t-values. $R^2 = 0.93$. D.W. = 1.64. Estimation Period: 1972-95.

This result stands very much in the tradition of the Austrian wage-formation process, showing almost full compensation for the inflation of the past year (which could also be interpreted as a simple expectation formation). The error correction term of the level equation (1) is significant and correctly signed.

2.2 The unemployment equation

Estimation of equation (3) gives the following result:

$$\Delta U = -0.153\Delta\left(y - y^{pot}\right) - 0.079\Delta\left(y - y^{pot}\right)_{-1} - 0.943\Delta\left(\Delta app + \Delta app_{-1}\right)/2$$

(-5.64) (-2.57) (-3.91)

$$+0.016\Delta nwlc + 0.076\Delta union + \varepsilon_{\Delta u} \qquad (3')$$

(3.62) (2.60)

Numbers in parentheses are t-values. $R^2 = 0.827$. D.W. = 2.47. Estimation Period: 1973-95.

This result conforms reasonably well to the expectations. Positive deviations of real output from its trend development reduces unemployment to a rather modest extent. As observed in many other studies, rising non-wage costs foster unemployment.

2.3 The price equation

According to equation (4), the evolution of prices is determined by the relation between wages and productivity and foreign prices. We get the following result:

$$\Delta p = 1.168 + 0.476\left(\Delta ulc + \Delta ulc_{-1} + \Delta ulc_{-2}\right)/3 + 0.080\Delta p^m + \varepsilon_{\Delta p} \qquad (4')$$

(2.131) (5.625) (1.79)

Numbers in parentheses are t-values. $R^2 = 0.754$. D.W. = 1.724. Estimation Period: 1973-95.

Thus, price developments are mainly determined by changes in unit labour costs.

2.4 The production function

The production function is, as mentioned, solved for productivity, and estimation of equation (5) exhibits the following:

$$y - 0.75h - 0.25k = 3.515 + 0.013T + 0.688\left(y - y^{pot}\right) + \varepsilon_y \qquad (5')$$

(772.17) (42.53) (5.11)

Numbers in parentheses are t-values. $R^2 = 0.988$. D.W. = 0.765. Estimation Period: 1970-92.

Accordingly, overall productivity is determined by an autonomous time trend and by the output gap.

3. Results of the system estimation and interpretation

A system approach has some advantages over a partial equilibrium approach when estimating the natural rate of unemployment. In particular, the labour market gap will not only have an affect on wage inflation, nor will the product market gap have an affect on price inflation alone. Instead, the parameters that define the natural rate of unemployment and potential output are themselves estimated jointly with the parameters in the wage and price equations. The specifications for the system are those of the single equation estimation and are repeated here for the sake of convenience:

$$w = \alpha_0 + \alpha_1 p^c + \alpha_2 q^{pot} + \varepsilon_\omega \qquad (6)$$

$$\Delta w = \beta_1 \Delta p^c + \beta_2 \Delta p^c_{-1} + \beta_3 \Delta\left(U - U^{nat}\right) + \beta_4 \Delta\Delta\left(U - U^{nat}\right) + \beta_5 \Delta\left(p^c - p\right) + \beta_6 \Delta\left(p^c - p\right)_{-1} + \beta_7 \varepsilon_{\omega-1} + \varepsilon_{\Delta\omega} \qquad (7)$$

$$\Delta U = \gamma_1 \Delta\left(y - y^{pot}\right) + \gamma_2 \Delta\left(y - y^{pot}\right)_{-1} + \gamma_3 \Delta\left(\Delta app + \Delta app_{-1}\right)/2 + \gamma_4 \Delta nwlc + \gamma_5 \Delta union + \varepsilon_{\Delta u} \qquad (8)$$

$$\Delta p = \delta_0 + \delta_1\left(\Delta ulc + \Delta ulc_{-1} + \Delta ulc_{-2}\right)/3 + \delta_2 \Delta p^m + \varepsilon_{\Delta p} \qquad (9)$$

$$y - 0.75h - 0.25k = \zeta_0 + \zeta_1 T + \zeta_2\left(y - y^{pot}\right) + \varepsilon_y \qquad (10)$$

where

$$U_t^{nat} = 1.42 + \gamma_3 \sum_{k=71}^{t}\left(\Delta app_k + \Delta app_{-1}\right)/2 + \gamma_4 \sum_{k=71}^{t} \Delta nwlc_k + \gamma_5 \sum_{k=71}^{t} \Delta union_k \qquad (11)$$

$$y^{pot} = 0.75h^{pot} + 0.25k^{pot} + \zeta_0 + \zeta_1 T \qquad (12)$$

$$q^{pot} = -0.25h^{pot} + 0.25k^{pot} + \zeta_0 + \zeta_1 T + l^{pot} \qquad (13)$$

$$\Delta y^{pot} - \Delta h^{pot} + \Delta l^{pot} \equiv \Delta q^{pot} = -0.25\Delta h^{pot} + 0.25\Delta k^{pot} + \zeta_1 + \Delta l^{pot} \qquad (14)$$

Compared with the single-equation estimates, no proxies for potential output or trend labour productivity plus the natural rate of unemployment are used. Instead it is necessary to substitute expressions (11)-(14) into equations (6)-(10).

The constant 1.42 in equation (11) is the unemployment rate in August 1971, which was the lowest rate during the sample period. The determinants of potential output have been expressed at their potential levels. For labour input (*h*), this is done by calculating total man-hours based on trend average hours per employee, and employment levels consistent with unemployment at its natural rate. The stock of capital is corrected by a capacity utilisation rate.[4] l is the log of the annual working hours per employee.

The system has been estimated using non-linear three-stage least squares. Endogenous variables in the system are the contemporaneous values of the dependent variables as well as the growth of labour productivity. The estimation period is from 1970 to 1995.

Table 1
System estimates

Equation				
w	ΔU	Δw	Δp	y**–0.75**h**–0.25**k
$\alpha_0 = -9.827$ (-2.51)	$\gamma_1 = -0.071$ (-2.48)	$\beta_1 = 1.192$ (6.59)	$\delta_0 = 2.253$ (5.11)	$\zeta_0 = 3.511$ (521.86)
$\alpha_1 = -1.328$ (17.07)	$\gamma_2 = -0.119$ (-5.12)	$\beta_2 = 0.486$ (2.28)	$\delta_1 = 0.380$ (5.22)	$\zeta_1 = 0.013$ (26.51)
$\alpha_2 = 0.930$ (2.83)	$\gamma_3 = 0.903$ (-2.94)	$\beta_3 = -2.986$ (-3.58)	$\delta_2 = 0.104$ (2.70)	$\zeta_2 = 0.469$ (5.31)
$\alpha_3 = -0.045$ (-3.669)	$\gamma_4 = 0.016$ (2.66)	$\beta_5 = -1.507$ (-7.38)		
	$\gamma_5 = 0.065$ (2.201)	$\beta_6 = -0.690$ (-2.92)		
		$\beta_7 = -0.153$ (-1.90)		
$R^2 = 0.997$	$R^2 = 0.543$	$R^2 = 0.810$	$R^2 = 0.666$	$R^2 = 0.991$
$\overline{R}^2 = 0.996$	$\overline{R}^2 = 0.421$	$\overline{R}^2 = 0.684$	$\overline{R}^2 = 0.627$	$\overline{R}^2 = 0.990$
SEE = 0.028	*SEE* = 0.311	*SEE* = 1.582	*SEE* = 1.116	*SEE* = 0.009

The system estimation results imply the following expressions for the natural rate of unemployment and potential output:

$$U_t^{nat} = 1.42 - 0.903 \sum_{k=71}^{t} \left(\Delta app_k + \Delta app_{-1}\right)/2 + 0.016 \sum_{k=71}^{t} \Delta nwlc_k + 0.065 \sum_{k=71}^{t} \Delta union_k \qquad (15)$$

$$y^{pot} = 0.75h^{pot} + 0.25k^{pot} + 3.511 + 0.013T \qquad (16)$$

Estimates of the historical path of potential output and output as well as the output gap and the growth rates of potential output are shown in Figures 1, 2 and 3.

One feature of the results is that the output gap fluctuated during the 1970s with spikes in 1974, 1977 and 1980. Subsequently the output gap decreased with a tendency to increase after 1988 (due to the lack of capital stock data, potential output could not be estimated for the years after 1992).

---

4 In a first approximation the August unemployment rate is considered to be the natural unemployment rate. In the yearly cycle the August unemployment rate is usually the lowest one. Total employees plus the fraction of the unemployed not considered as structural is the potential supply of employees (E^{pot}). Potential hours per employee (L^{pot}) is calculated as the long run equilibrium of an OLS regression of hours per employee on a constant, on trend and lagged hours per employee: ($L=f(1,T,L_{-1})$). Potential labour input is hence $H^{pot}=L^{pot}\,E^{pot}$. A capacity utilisation index is constructed via yearly survey results of the "Konjunkturtest" (KT) and the "Investitionstest" (IT). An OLS regression of (100-IT) on a constant and KT yields the capacity utilisation variable with which the private capital stock is adjusted to yield potential private capital. Public sector real capital is assumed to be fully employed.

Figure 1

Output and potential output (in 1,000 billions of Austrian schillings)

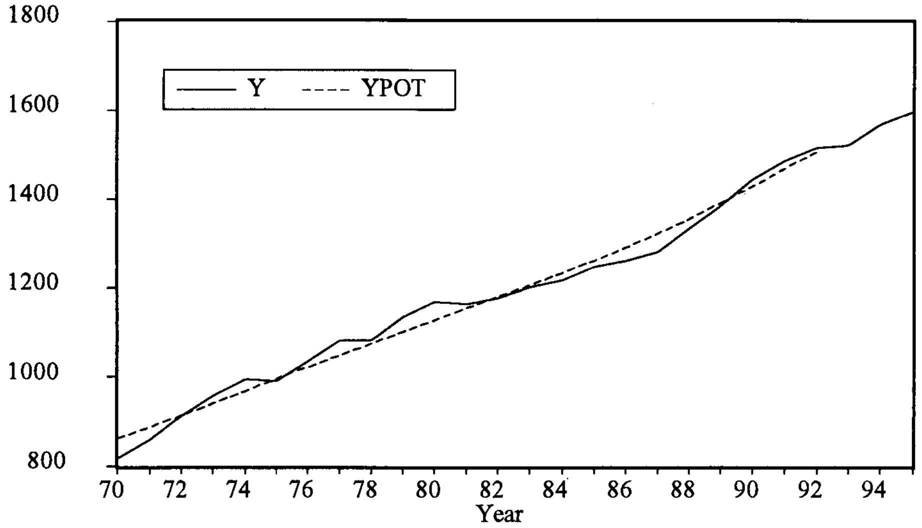


Figure 2

Output gap (as a percentage of potential output)

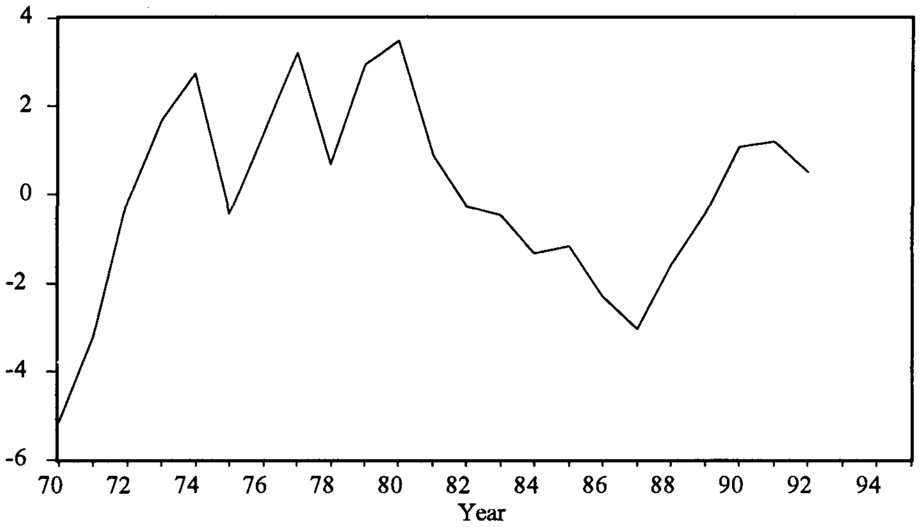
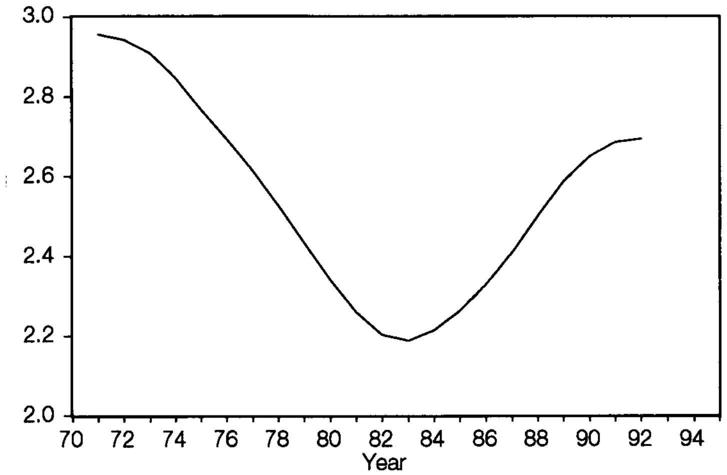


Figure 3

Potential output (annual percentage changes)



Estimates of potential output along with the contributions to growth of changes in potential factor inputs and multifactor productivity are shown in Table 2.

Table 2
Contributions to the growth of potential output
Annual percentage changes

	1970-80	1981-88	1989-92
Potential output	2.82	2.19	2.09
Labour	0.41	0.29	0.53
Capital	0.86	0.58	0.50
Multifactor productivity	1.55	1.32	1.06

Slower growth in factor inputs contributed most to the slowdown in the growth of potential output during 1981-88. In the period 1989-92 labour growth picked up whereas capital did not return to its contribution of the 1970s. Multifactor productivity gradually decreased over the whole estimation period.

The natural rate of unemployment is estimated to have increased steadily from a low of 1.5% in 1971 to a high of 6.3% in 1995. The increase in the natural rate of unemployment gained momentum particularly during the 1980s. The unemployment gap widened during the 1980s but narrowed in the last couple of years.

Mirroring the development of potential output, the same three periods seem to have crystallised, with different trends in actual and natural unemployment – pretty stable development during 1970-80, a drastic pickup after 1980 with a narrowing of the unemployment gap after 1988.

According to the estimation, the bulk of the increase in unemployment after 1980 can be attributed to a rising natural unemployment rate.

Figure 4
Actual and natural unemployment rates

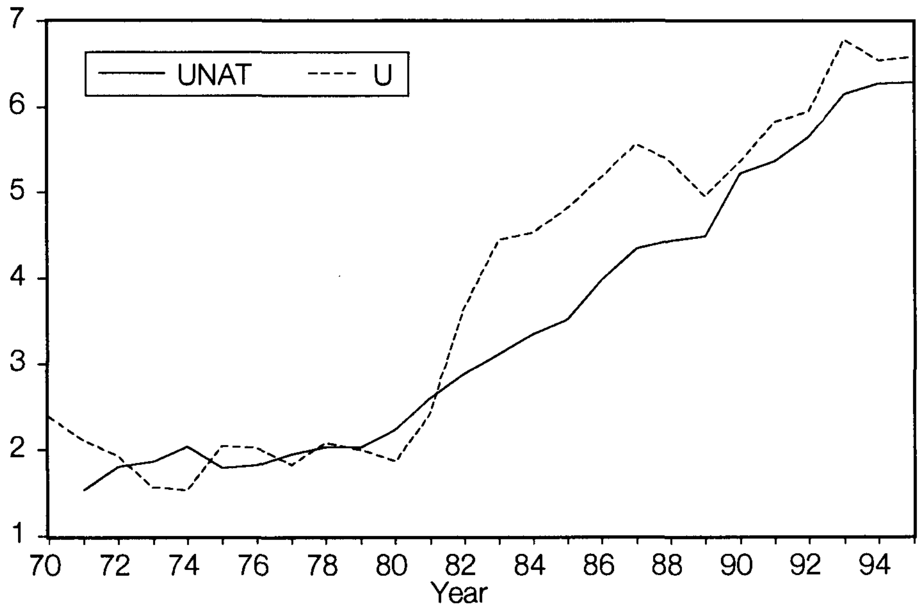


Table 3
Contributions to changes in the natural rate of unemployment

	1970-80	1981-88	1988-95
Changes in actual unemployment rate	-0.2	3.0	1.6
Changes in natural rate	0.7	1.8	1.8
Due to non wage labour cost	1.7	0.7	0.7
Due to changes in apprentices	-1.2	1.2	1.5
Due to changes in union density	0.2	-0.1	-0.4

Conclusions

What lesson may we draw from this exercise? The rise in unemployment after 1980 seems to be determined by structural factors which raised the natural unemployment rate. We estimated a steady increase of the NAIRU from 1.5% in 1971 to 6.3% in 1995. The rise gained momentum particularly in the 1980s and the gap between the NAIRU and actual unemployment widened in this period, but has narrowed again recently. From this point of view, measures to remove the structural impediments responsible for the malfunctioning of the labour market are called for. We have identified three microeconomic factors, namely non-wage labour cost union density and the share of apprentices in the labour force.

One policy conclusion from this result would be that measures should be taken to speed up an easier integration of young people into the Austrian labour market and the flexibility of the labour market should be enhanced – for example with respect to aspects of labour market regulation which primarily affect the cost of labour; e.g. labour cost advantages of employing part-time workers, cost of overtime pay and so on. Though we regard our results for the NAIRU as plausible, it has to be pointed out that they seem to be very sensitive to parameter changes and to the specification of the underlying equations and assumptions. Thus, policy conclusions should be drawn with great care.

What the results certainly dismiss is that fiscal or monetary policies have a tight link with the rise of the unemployment rate in Austria.

Appendix: Data sources

In the empirical analysis, annual data were used, ranging from the beginning of the 1970s until 1995 (depending on availability). Most time series are taken from Österreichisches Statistisches Zentralamt: Österreichs Volkseinkommen, as provided by WSR Data Bank.

References

Adams, Ch. and D. T. Coe (1990): "A Systems Approach to Estimating the Natural Rate of Unemployment and Potential Output for the United States". IMF *Staff Papers*, No. 37, pp. 232-293.

Coe, D.T. and T. Krueger (1990): "Why is Unemployment So High at Full Capacity? The Persistence of Unemployment, the Natural Rate, and Potential Output in the Federal Republic of Germany". IMF *Working Paper*, No. 90/101.

Ball, L. (1996): "Disinflation and the NAIRU". NBER *Working Paper*, No. 5520.

Gartner, Ch. (1995): "Potential Output and Output Gap". Oesterreichische Nationalbank, *Reports and Summaries*, No. 3, pp. 41-50.

Hahn, F. R. and G. Rünstler (1996): "Potential-Output-Messung für Österreich". Österreichisches Institut für Wirtschaftsforschung, *Monatsberichte*, No. 3, Vienna.

Handler, H. (1989): "Grundlagen der österreichischen Hartwährungspolitik". Vienna.

OECD (1994): *The OECD Jobs Study*. Paris.

Pichelmann, K. (1989): "Unemployment Dynamics, Wage Flexibility and the NAIRU in Austria". Institute for Advanced Studies, *Research Memorandum*, No. 258, Vienna.

Pichelmann, K. and A.-U. Schuh (1996): "The NAIRU-Concept – A Few Remarks". Institute for Advanced Studies, *Economic Series*, No. 36, Vienna.

Wehinger, G. (1990): "Potential Output und strukturelle Arbeitslosigkeit in Österreich". Institute for Advanced Studies, Vienna.

Wörgötter, A. (1983): "A Note on the Stable Phillips Curve in Austria". *Empirica*, No. 10(1), pp. 29-40.

Comments on: "A systems approach to the determination of NAIRU, inflation and potential output in Austria" by Friedrich Fritzer and Heinz Glück

by Palle Andersen

As the title suggests, this paper follows the approach of Coe and Adams (1990) and Coe and Krueger (1990) for the United States and Germany. This is certainly the right approach to these issues; yet I have some problems with the specifications used by Fritzer and Glück, their estimates and the final results for the NAIRU. Although I cannot quite find the source, I also think there is some inherent instability in the model.

I do not have much say about their review of the literature for Austria, except that the authors have not received much help from this source. Perhaps one estimate could have been updated and used and that is Pichelmann's finding that distributional issues have forced up the rate of unemployment in Austria. Since distributional issues are usually not included in wage negotiations but left to fiscal policies, it is quite possible that in a period when fiscal consolidation has made it more difficult to apply fiscal policy to solve distributional issues, these may have contributed to the rising unemployment. In other words, distributional proxies could have been tested in the equation for unemployment.

Let me then turn to the specification of the model and the individual equations. The basic wage equation is based on the target real wage bargaining model, with real consumption wages targeting potential productivity, and estimated as an error correction model. This differs from the standard expectation-augmented Phillips curve in that deviations between actual and equilibrium unemployment do not necessarily result in a continuous deceleration or acceleration of nominal wage changes. Given the close link between Austria and Germany, it is also worth noting that Fritzer and Glück are able to confirm that the wage equation found by Coe and Krueger for Germany also applies to Austria. Moreover, as for Germany, changes in the terms of trade (modelled as the difference between changes in respectively output and consumer prices) significantly affect nominal wages. In fact, in their estimates, the overall coefficient on consumer prices is actually 0 while output prices obtain a coefficient of 1.7 which seems a bit on the "high side" to satisfy the homogeneity condition. Perhaps, the authors should have used output prices instead of consumer prices in their target equations as the "story" of the error correction equation is clearly one of nominal wages following output prices.

The unemployment equation is specified in first differences which is sensible since the rate of unemployment follows an A2-process for the period considered. In other words, there is almost complete hysteresis. Including changes in non-wage labour costs, unionisation and the proportion of apprentices as explanatory variables is also sensible, whereas I am not quite sure why the authors prefer *changes* in the output gap; after all, by construction, the output gap is a stationary variable.

While the wage equation "overshoots" the homogeneity condition by assigning a very high coefficient to price changes, we have the reverse problem in the price equation as the sum of the nominal coefficients is less than 0.6. This clearly creates some short-run instability in the model which, undoubtedly, would show up as very large short-run swings in the implicit profit share. Given that Austria is a small and open economy, I was also surprised to see such a small impact of foreign prices. Perhaps more lags should have been included or possibly a joint term in import prices and the output gap as the pass-through of import prices might be influenced by cyclical conditions in Austria.

In the production function, which is specified with total factor productivity as the dependent variable, I was slightly disturbed by the rather high coefficient on the ratio of actual to potential output. This is a very strong Verdoorn-effect and really implies that by running the economy above potential one can generate a permanent rise in the level of total factor productivity; I suspect that this might be another source of the instability referred to above[1].

There are relatively few major changes in going from the single-equation estimates to the system approach and it is reassuring to see a confirmation of the importance of output prices in the setting of wages and of the high degree of real wage flexibility in Austria. I am, however, a bit disturbed by the further drop in the nominal coefficients of the price equation and by the implicit rise in the short-run instability of profits.

Finally, let me turn to the implications of the estimates for the natural rate of unemployment as presented in Figure 4 and Table 3. The major cause of the rise in the NAIRU since 1980 seems to be the fall in the number apprentices as a proportion of the labour force. That is a very strong result and rather unique among European countries and I wonder what caused it and what could be done to reverse the trend. Is it possible that this variable could be acting as a proxy for something else and something more fundamentally disturbing? In this context, I also wonder how one can be so certain that the rise in the natural rate had nothing to do with fiscal and monetary policy? I certainly hope that this is true, but since the Austrian unemployment story is essentially a hysteresis story, I would not give a "not guilty" verdict to policies without further proof.

---

1 Incidentally, let me note that it would be helpful to know how the variables have been defined and measured and what sources have been used. This would also enable other readers to reproduce and test the results.

An alternative to the mainstream model of inflation with an application to Switzerland

Franz Ettlin

The inflation rate and the question of its determining process is not the only concern for central bank policy, but in most of the developed countries undoubtedly a primary one. The model of price change underlying the thinking of economists in general, and of central bankers as well as other policy makers in particular, is dominated by the so-called price Phillips-curve approach. It can be characterised by three main features. Firstly, the inflation rate is crucially dependent by the degree of aggregate resource utilisation. A couple of statistical indicators are preponderantly used, either alternatively or in conjunction, to provide representative utilisation measures. Most commonly applied is the general unemployment rate, as an inverse measure, and the GDP-gap, the difference between the economy's actual and potential real output, as a direct measure. To avoid obvious repetitive-type formulations, the unemployment rate will in general be singled out in the following text as the representative measure of aggregate resource utilisation. Secondly, the inflationary impulse generated by the degree of aggregate resource utilisation is partially or fully propagated forward in time by inflationary expectations which are formed on the actual inflation process. With partial propagation, this model can be summarised by a negatively sloped relationship between inflation and the unemployment rate. Thirdly, in case inflationary expectations fully respond to actual inflation, there is a unique level of the unemployment rate at which the inflation rate is constant. This is the concept of the so-called NAIRU, the non-accelerating-inflation rate of unemployment. Any given positive or negative deviation of unemployment from this unique level of the NAIRU is supposed to generate a corresponding acceleration or deceleration of the inflation rate. This implies that the long-run Phillips curve forms a vertical line in a two-dimensional diagram with inflation on the vertical and unemployment on the horizontal axis. Looking at the international research output on inflation and related economic policy issues, this NAIRU view of the inflation process is taken seriously by many economic theorists, econometricians and policy practitioners, not least in the central banks and some prestigious multinational economic and monetary organisations.

It is the position taken in this paper that the Phillips curve and the associated NAIRU are seriously flawed concepts, and that a very adequate alternative explanation of the inflation process exists. This position is based both on theoretical and empirical grounds.

A first argument against modelling the inflation process according to the Phillips curve derives from statistical considerations of the stationarity properties of the time series for the inflation rate, on the one hand, and the unemployment rate (or the GDP-gap based on an aggregate production function and potential input approach) on the other. There is evidence for Switzerland, as well as for other countries, that the inflation rate preferably should be viewed as a stationary I(0) process and the unemployment rate as a non-stationary I(1) process. This implies that for a valid model the inflation rate should in principle be related to the change in the unemployment rate, as they are both stationary I(0) variables. In other words, a stationary relationship involving the non-stationary level of the unemployment rate could only exist by cointegration with the price level, not its rate of change.

A second argument against modelling inflation according to the NAIRU approach is based on the solution of the second-order differential equation corresponding to the standard Phillips-curve model with NAIRU properties.

That model forming a negative relationship between the rate of change of the inflation rate and the level of unemployment can be stated as:

$$\ddot{p}(t) = -\alpha u(t)$$

where the symbol $\ddot{p}(t)$ stands for the second time derivative of the price level p at time t and $u(t)$ for the corresponding unemployment rate (measured as a deviation from an inflation-neutral base level). From this the following general solution for the price level can be derived:

$$p(t) = c_1 + c_2 t + \alpha \int_{t_0}^{t} u(\tau)\tau \, d\tau - \alpha t \int_{t_0}^{t} u(\tau) d\tau$$

with initial conditions for the price level and the inflation rate and $p(t_0) = p_0$ and $\dot{p}(t_0) = \dot{p}_0$ respectively.

From an economic point of view, the price level theory implicit in this deterministic integral equation appears rather implausible. Apart from the homogeneous solution which forms a constant-coefficient linear time trend of unspecified origin, the price level is determined only by the complicated time-weighted history of the unemployment rate. The farther back in time a given unemployment rate was experienced, the stronger is its cumulated influence on the price level.[1] There are no further determinants. In particular, any influence from the uneven temporal development of prices in the rest of the world as well as of exchange rates remains completely neglected in this pure NAIRU version of the price Phillips curve.

There are, of course, also less extreme versions of Phillips type relations, including A.W.H. Phillips' famous original paper, which make some allowance, at least in principle, for a secondary influence from import price changes in the inflation process. It is, however, the contention of this paper – in broad agreement with the so-called Scandinavian model of inflation[2] – that in the typical open economy the domestic currency price of foreign goods and services is not only a secondary, but a primary force in domestic inflation. This important determinant can, in turn, be split into foreign price changes measured in foreign currency and changes in the domestic price of foreign currency. The neglect of this aspect, or its only secondary consideration in several inflation models, seems to be representative of a current norm among economists in general.

Cyclical fluctuations in domestic employment and output play also a significant role in the inflation process, but their direct influence tends to be much smaller on average than the influence from changes in foreign prices and the exchange rate. Moreover, as suggested in the first critical argument above, in many, if not most countries, this secondary force on inflation is likely to be related to the changes and not the level of the unemployment rate.[3]

---

1 This shows up still more clearly in the discrete-time equivalent to the integral solution equation: $p_t = c_1 + c_1 t - \alpha \sum_{\tau=0}^{t-1} (1+\tau) u_{t-\tau}$.

2 For a survey in English of the Scandinavian input-output model of inflation by its Norwegian originator see Aukrust (1997). The first econometric implementation of the corresponding (also noneconometric) study for Sweden by Edgren, Faxén and Odhner (1973) was Ettlin (1974). It was integrated into a full econometric model context in Ettlin et al. (1979).

3 In Ettlin (1974) and Ettlin et al. (1979) the wage and price levels, not their respective rates of change, were furthermore explicitly related to the level of the unemployment rate. This implied an aggregate domestic supply curve with a positive slope with respect to the employment rate rather than a Phillips-curve relation. The former type of relationship regarding wages has recently been termed "the wage curve" by Blanchflower and Oswald (1994), who found it empirically confirmed in a huge set of disaggregated data from a large number of countries. Perhaps, "the price curve" would be an appropriate term for the price level relationship with the unemployment level. Sargan (1964) seems to be the first econometric study of price and wage formation which made allowance for the primary

A basic version of the alternative inflation model for an open economy can be derived along the following lines. A country's gross output is assumed to be generated according to two distinct log-linear aggregate production functions, one for the tradables and the other for the non-tradables sector (called exposed and sheltered sectors respectively in the Scandinavian model literature). Labour, capital as well as imports serve as production factors, and technical progress proceeds at differing rates in the two sectors. Demand for the output of domestically produced tradables and non-tradables depends log-linearly on domestic and foreign real income and on the supply prices of domestic and foreign tradables and non-tradables. In this type of environment, instantaneous profit maximisation with regard to all the factor inputs yields long-run optimal prices of tradables and non-tradables. The logarithm of the prices of domestically produced tradables and non-tradables is a log-linear function of the factor prices for labour, capital and imports with the coefficients determined chiefly, and in case of constant returns to scale entirely, by the respective output elasticities of the factors.

Money wage rates generally respond to the conditions pertaining in the sector exposed to international competition. This, firstly, refers to the development of foreign prices measured in domestic currency and, secondly, to the faster rate of technical change in the largely goods-producing tradables sector. It implies that, on average, unit labour costs and prices rise faster for non-tradables than for tradables. As a result, the total rate of price change of consumer expenditures, with their large share of non-tradables will tend to exceed the rate of price change of tradables.

The latter point provides an important part of the explanation for the persistence of above-target rates of change of consumer prices even in countries like Germany and Switzerland. There, in spite of monetary policies mainly geared to price stability, the target rates of inflation of zero or one percent were overshot by at least a couple of percentage points on average since the mid-1970s. If the annual rate of growth of technical progress for non-tradables is, for example, 3 percentage points lower than for tradables, then the price rise of non-tradables will tend to exceed that of tradables by 3 percentage points. With a share of, say, two-thirds for non-tradables in total consumer expenditures a 2% inflation rate would result on that account alone.[4]

The specification of the alternative inflation model is, like the price Phillips-curve, a quasi-reduced form or system's solution of structural equations for wages and prices. But as outlined above, the structural equations of the alternative model are quite distinct from the corresponding Phillips-curve relations. The result is a model specification in which the inflation rate of consumer prices (ΔLPC) depends on the current and particularly also the lagged rate of change of import prices ($\Delta LPIM$) as well as a positive constant for the intercept reflecting both the effect of the sectoral productivity growth differential on total unit labour costs and the average rate of growth of unit capital costs and unit indirect taxes.

The relevant price of tradables is, in fact, more directly related to the price of exports. But the latter is an endogenous variable that is substituted out in the reduced-form specification in favour of its major determinant, the price of imports, which has a more exogenous character.

The influence of the domestic business cycle on wage and price determination is reflected also by the (lagged) unemployment rate as in Phillips-curve models, but, in contrast to the latter, it is a priori expected that the change rather than level of the (logarithm of the) unemployment rate ($\Delta LUNR$) affects the inflation rate. The short-run elementary version of the alternative model can thus be formally stated as :

---

importance of import prices and correctly interpreted the role of the unemployment variable, at least in the long-run solution he derived from the estimated equations. But, as a 90% adjustment to a new equilibrium was estimated by Sargan to take almost 11 years, his results did not differ that much in the short and medium run from conventional Phillips-type studies.

4 On that issue, see also Krugman (1996).

$$\Delta LPC = \alpha_0 + \alpha_1 \Delta LPIM + \alpha_2 \sum \gamma_i \Delta LPIM_{-x-i} - \alpha_3 \Delta LUNR_{-y}$$

The estimated price change equations presented in Table 1 provide empirical evidence on these issues for Switzerland. Since the change of the logarithm of consumer and import prices as well as the change in the logarithm of the unemployment rate are stationary I(0) variables, a stationary short-run relationship is estimated in equations 1 and 2. They are simplified versions of the alternative error-correction model to be shown in Table 3. To limit the specification at first to the most essential variables only, the sample period in Tables 1 and 2 ends in 1994.[5]

Table 1

Basic quarterly rate of change equations of the Swiss consumer price index 1976Q1-1994Q4

Dependent variable: Δ *LPC*				
Regressors:	**Equation 1**	**Equation 2**	**Equation 3**	
Intercept	0.006 (10.21)	0.007 (11.16)	0.004 (3.03)	Regressors of alternative model
$\Delta\ LPIM$	0.200 (6.70)	0.207 (7.29)	0.187 (4.69)	
$\Delta\ M8PIM_{-3}$	0.340 (6.19)	0.308 (5.80)	0.291 (4.01)	
$\Delta\ LUNR_{-6}$		-0.008 (3.00)	-0.011 (2.32)	
$LUNR_{-6}$			0.001 (1.37)	Regressors of mainstream Phillips-curve model
$\Delta\ LPC_{+1}$			0.080 (0.79)	
$\Delta\ LPC_{-1}$			0.101 (0.97)	
Standard t-statistics are shown in parenthesis below the estimated coefficients.				
Summary statistics:				
Standard error of regression	0.005 (0.48%)	0.005 (0.45%)	0.005 (0.46%)	
Adjusted R-squared	0.446	0.501	0.494	
Durbin Watson statistic	1.741	1.799	2.050	
Number of observations	76	76	76	

5 The variables are explained at the end of this paper.

In equation 1 the change of the logarithm of import prices appears first unlagged with a significant coefficient of 0.20. This reflects mainly the direct input price effect from imports on consumer prices. The second import price term with a similarly significant coefficient of 0.34, refers to the change in an eight-quarter moving average, which is lagged by three quarters. This influence on consumer prices is to be interpreted as an indirect one via the determination of wages and unit labour costs. Finally, the highly significant intercept term of 0.006 implies a time trend of 0.6% per quarter or 2.4% per annum in the price level. According to the alternative model, it is in principle agreement with the expected result from lower productivity growth in non-tradables than in tradables production. This interpretation is supported by the fact that the average inflation rate for consumer services is much higher than for consumer goods. A part of this time trend is, however, also due to the average rate of growth of nominal unit capital costs and unit indirect taxes, both of which do not appear explicitly as regressors.

Equation 1 explains 45% of the variance of the inflation rate without making use of the positively autocorrelated characteristic of inflation. Equation 2 contains additionally the lagged change in the unemployment rate, also in accordance with the alternative model. The specified time delay of six quarters for the change of unemployment reflects mainly the lag in the determination of wages and unit labour costs but also some lag of consumer prices behind unit labour costs. This time delay approximately matches the total mean delay of the lagged moving average import price term identified above as the primary empirical determinant of consumer prices via wages and unit labour costs. The correctly signed and significant coefficient estimate of the change of the logarithm of the unemployment rate is -0.008. This implies, for example, that a doubling of unemployment from a level of 0.5% to 1% or from 1% to 2% will, after a delay of six quarters, reduce the inflation rate by about 0.6%. But with no further change in the level of unemployment assumed, the inflation rate will thereafter be unaffected by the unemployment situation. In other words, in a two-dimensional Phillips-curve diagram, with the inflation rate on the vertical and the unemployment rate on the horizontal axis, this long-run partial relationship forms a horizontal line,[6] in contrast to a vertical line for the NAIRU model. The height of this horizontal line is, of course, mainly determined by the current and lagged changes in import prices and the factors subsumed into the intercept term. The comparison of equation 2 with equation 1 shows that the addition of the change of unemployment variable, although significant, affects the estimated coefficients of the other regressors only to a limited extent. According to the adjusted R-squared statistic the explained portion of the variance of the inflation rate rises, but only moderately from about 0.45 to 0.50. In other words, the results of equations 1 and 2 suggest that over an almost twenty-year period the actual variations in the unemployment rate within a range of almost 0–5% were only a very secondary force in Switzerland's inflation experience compared to the dominant forces that emanated from variations in import prices and a positive trend factor, presumed to reflect the productivity growth differential between the tradables and non-tradables sector and the growth of nominal unit capital costs and unit indirect taxes.

In Equation 3 a basic set of regressors from the mainstream Phillips-curve model is added to the elementary version of the alternative model. All the coefficients of the alternative model variables remain correctly signed and significant, whereas no coefficient of the mainstream model variables achieves standard levels of significance. The coefficient of the unemployment level is wrongly signed, and price change expectations modelled generously on the actual inflation rates one period ahead (!) and one period behind have only a minor impact on the coefficients of the alternative model. Moreover, the rather small size of the coefficients of the actual inflation rates surrounding the current period suggests that inflation expectations do not play the crucial role attributed to them in the mainstream model. It is clear from these empirical results for Switzerland that the alternative model entirely encompasses the mainstream model.

---

6 Positive and negative changes of unemployment will, cet. par., lead to temporary deviations of the inflation rate from this horizontal line. These transitory observations will result in a scatter with a negative slope like a short-run Phillips curve.

Table 2 shows some estimation results for the mainstream model alone. In equation 4 the level of the unemployment rate is entered for two successive periods, with the six quarter lag again being empirically determined. According to the mainstream model the respective coefficients should either be significantly negative or one significantly negative and the other insignificantly different from zero. This is clearly not born out by the results which show a negative coefficient an lagged unemployment being matched by a similar positive coefficient an unemployment lagged by an additional period. This points towards the change of unemployment effect on inflation in accordance with the alternative model. With these kinds of results being rather typical for other countries as well, it comes as no real surprise that the dedicated pursuit of the Phillips-curve approach has, for various countries, resulted in the hypothesis of a natural rate of unemployment which empirically tends to follow the actual rate: The presumed unemployment level, measured as a deviation from the adapting natural rate, becomes, in fact, an approximation to the change of unemployment!

Table 2

Mainstream-related quarterly rate of change equations of the Swiss consumer price index 1976Q1-1994Q4

Dependent variable: ΔLPC			
Regressors:	**Equation 4**	**Equation 5**	**Equation 6**
Intercept	0.002 (1.88)		0.004 (3.99)
$LUNR_{-6}$	-0.006 (1.97)	-0.005 (1.52)	
$LUNR_{-7}$	0.005 (1.75)	0.004 (1.25)	
ΔLPC_{+1}	0.343 (3.56)	0.439 (5.28)	
ΔLPC_{-1}	0.348 (3.48)	0.443 (5.05)	0.451 (4.29)
Standard t-statistics are shown in parenthesis below the estimated coefficients.			
Summary statistics:			
Standard error of regression	0.005 (0.52%)	0.005 (0.53%)	0.006 (0.58%)
Adjusted R-squared	0.340	0.317	0.188
Durbin-Watson statistic	2.446	2.677	1.950
Number of observations	76	76	76

The future and past actual inflation rates in equation 4 have correctly signed significant coefficients summing to 0.69. But this result only appears because the estimated equation is primarily a statistical lead-lag autocorrelation relationship with little autonomous economic substance added by the presumed impulse variable, the unemployment level.

In equation 5, where the intercept term is suppressed, the coefficients of the two unemployment level variables again show opposite signs, which accords with the alternative model's change of unemployment specification, although these coefficients are not significantly different from zero according to standard criteria. This again leaves it foremost to the two lead-lag actual inflation rates to explain the variance of the current inflation rate. Under these circumstances and in view of the univariate autocorrelation relationship, the correct signs, the magnitude and the significant t-values of the latter two coefficients come as no surprise. But this is not sufficient to confirm the crucial propagator role of inflation expectations as expounded in the mainstream view. Finally, the first-degree positive autocorrelation of the time series for inflation is illustrated in equation 6 without interference from any other variables. This univariate relationship, again, is no proof that inflation is preponderantly or fully propagated by inflation expectations.

Table 3

Error correction model for the consumer price index of Switzerland 1976Q1-1996Q4

Dependent variable: LPC	**Equation 7**	**Equation 8**
Regressors:	**Level equation**	**First-difference equation**
Intercept	3.648	
	(53.42)	
LPC_{-1}		0.180
		(2.07)
LPIM	0.039	0.074
	(2.90)	(1.80)
$LPIM_{-1}$	0.039	
	(2.90)	
LPIMERS	0.206	0.152
	(6.21)	(3.67)
$M8LPIM_{-3}$	0.322	0.246
	(27.29)	(5.28)
TR	0.006	0.005
	(61.82)	(7.81)
TR90Q392	0.005	0.002
	(22.68)	(1.87)
$M2LUNR_{-5}$	-0.009	-0.008
	(9.12)	(3.38)
VAT	0.008	0.011
	(2.61)	(2.88)
$LEVEL\text{-}EQUATION\ RESIDUAL_{-1}$		-0.571
		(5.68)

The numbers in parenthesis below the estimated coefficients are the absolute values of the t-statistic. Standard distribution assumptions concerning these t-statistics, however, apply only to the first-difference equation.

Summary statistics:

Standard error of regression		0.004 (0.36%)
Adjusted R-squared		0.685
Durbin-Watson statistics	0.970	1.858
F-statistic (8,75)		23.526
Number of observations	85	84

Equation 2 shown in Table 1 above represents a basic version of the alternative inflation model. Equations 7 and 8 in Table 3 below provide a fuller version which makes use of the cointegration and error correction approach. Equation 7 estimates the relationship between the consumer price level and the level of current and lagged import prices, a segmented time trend, and the lagged level of the unemployment rate. Finally, the model explicitly incorporates the changeover from the sales tax to the value-added tax in 1995, which was by far the largest discretionary increase in the level of indirect taxation on consumption in the sample period.

The estimation results of the cointegrated level equation 7 confirm the a priori expectations of the alternative model to a remarkable extent. In particular, the magnitudes of the coefficients for the direct and the indirect influence of import prices on consumer prices seem to be relatively close to the actual share of import and labour costs in total consumption. Their sum of 61% would leave 39% to unit capital costs and unit indirect taxes, if the aggregate production functions incorporate constant returns to scale. The time trend reflects the lower productivity growth in non-tradables than in tradables production alluded to above and the trend rise of unit capital costs and unit indirect taxes (apart from the introduction of VAT, which is modelled explicitly). The second time trend variable reflects an intermittent phase of faster growth of productivity in tradables production in the early 1990s, which increased the inflation-contributing productivity differential with regard to non-tradables. The origin of this development was more rapid technological and structural change in export and import-competing activities, firstly in the wake of political and economic transformation in Eastern Europe, epitomised by German unification in mid-1990, and secondly by a speed-up of globalisation of economic activity in general. It is assumed in the specification, that the previous lower productivity growth differential was re-established in 1993 on account of a corresponding speed-up of productivity growth also in non-tradables. The level of unemployment with an average time lag of 5½ quarters reflects the pressure of demand in the labour market. It mainly affects price determination via wage formation. It should be observed that in Switzerland wages are mostly adjusted at the beginning of each year (few staggered contracts) with wage negotiations typically having taken place in the third and fourth quarter of the preceding year. This alone implies a mean lag for wage compensation of at least 3½ quarters behind observed import price developments and unemployment. Furthermore, there is some mean lag in the propagation of unit labour costs to consumer prices. Therefore, the empirically determined mean lags of 5½ quarters implied in the moving average terms for import prices and the unemployment rate appear quite plausible. Finally, the coefficient of the VAT dummy variable provides a reasonable estimate of the price level impact of the discretionary indirect tax increase in 1995.

The estimation results in Table 3 of the error-correction equation 8 for the quarter-to-quarter inflation rate confirm the estimates in the cointegration level equation 7. The highly significant coefficient of -0.57 for the lagged residual from the level equation – the error correction term – provides confirming evidence for the stationarity of the cointegration equation (which contains non-stationary I(1) variables). According to the magnitude of this coefficient more than half of any deviation of the actual price level from its equilibrium as determined by equation 7 will be corrected with a time delay of one quarter. After three quarters, more than 90% of the original deviation will have disappeared already. Moreover, as most of the other coefficient estimates of the first-difference equation 8 are quite similar to the corresponding estimates in the cointegrated level equation 7, the level equation residuals will tend to be small. The model accounts for 68% of the total variance of the inflation rate. The residual standard error of the first difference equation corresponds to 0.36% of the price level, which is less than half of the total standard deviation of the actual inflation rate.

Summarising, it is evident that the alternative inflation model outlined and estimated for Switzerland decidedly encompasses the mainstream Phillips-curve model. What does the estimated model in Table 3 imply with regard to the Swiss inflation experience during the last two decades? It was overwhelmingly related to the fluctuations in the rate of change of import prices expressed in domestic currency. Thus, the low inflation rates around 1976-78, 1986-88, and 1994-96 coincided with small or even negative current and lagged import price changes. Similarly, the high inflation episode around 1981-82 and, to a somewhat lesser extent, also the one around 1990-91, was accompanied by a large acceleration of import price changes. But the contribution of import prices to the average rate of inflation was only secondary compared with the much stronger time-trend related effect which is mainly attributable to the productivity growth differential. Finally, the development of the unemployment rate had only some secondary impact on the cyclical course of inflation.

The major implication for economic policy is, in brief, the suggestion that the widespread emphasis on the primary role of unemployment in the inflation process in general and the NAIRU view of a crucial threshold unemployment rate in particular does not appear to be justified. The derivation of detailed monetary policy implications requires that the alternative inflation model is analysed jointly at least with appropriate models of the exchange rate, interest rates and real economic activity. Although suitable empirical versions of these model requirements are presently available in Ettlin and Bernegger (1994) and Ettlin (1994, 1996), such a joint analysis, including relaxed exegoneity issues, will have to be the subject of a separate study.

List of variables

LPC, LPC_{+1}, LPC_{-1}	Natural logarithm of the consumer price index of Switzerland, for the current, succeeding and preceding quarter, respectively.
$LPIM$	Natural logarithm of the implicit deflator of total imports of goods and services.
$LPIMERS$	Natural logarithm of the unit value index of imports of energy, raw materials and semi-manufactures.
$M8LPIM_{-3}$	Eight-quarter moving average of *LPIM*, lagged three quarters.
$LUNR_{-6}$	Natural logarithm of registered unemployment as a percentage of the labour force, lagged six quarters.
TR	Integer linear time trend with 1976Q1=1.
$TR90Q392$	Integer linear time trend starting with 1990Q3=1 and ending in 1992Q4.
VAT	Dummy variable equal to zero until 1994Q4 and unity from 1995Q1 on, when a value-added tax replaced the previous more narrowly based sales tax.
Δ	First (backward) difference of the indicated variable.

References

Aukrust, Odd (1977): "Inflation in the Open Economy: A Norwegian Model", in Lawrence B. Krause and Walter S. Salant (eds.), *Worldwide Inflation*, The Brookings Institution, Washington, D.C., pp. 107-166.

Blanchflower, David and Andrew Oswald (1994): *The Wage Curve*, The MIT Press, Cambridge, Massachusetts and London, England.

Edgren, Gösta, Faxén and Odhner (1973): *Wage Formation and the Economy*, Allen and Unwin, London.

Ettlin, Franz (1974): "The International Dependence of Swedish Wages and Prices", in Franz Ettlin and Johan Lybeck, *Some Main Features of the STEP1 Econometric Model of Sweden*, presented at the Meeting of Project LINK at the Federal Reserve Board, Washington D.C., Econometric Project Working Paper, No. 76, EFI–Stockholm School of Economics.

Ettlin, Franz et al. (1979): *The STEP1 Quarterly Econometric Model of Sweden*, EFI–Stockholm School of Economics.

Ettlin, Franz (1994): "Economic Activity Effects of Monetary Policy through the Interest Rate and Exchange Rate Channels". Paper presented at the *International Symposium on Economic Modelling at the World Bank, Washington D.C.*, Swiss National Bank, Zürich.

Ettlin, Franz and Michael Bernegger (1994): "Interest rate determination across the maturity spectrum in a small open economy: the case of Switzerland", in Homa M. Scobie (ed.), *The European Single Market,* Chapman & Hall, London, pp. 45-64.

Ettlin, Franz (1996): "On the fundamental determinants of the Swiss franc exchange rate for the D-mark", in *The determination of long-term interest rates and exchange rates and the role of expectations*, Bank for International Settlements, Basle, pp. 18-27.

Krugman, Paul (1996): "Stable Prices and Fast Growth: Just Say No". *The Economist*, 31st August, pp. 17-20.

Sargan, J. Dennis (1964):"Wages and Prices in the United Kingdom: a study in econometric methodology", in P. E. Hart, G. Mills and J. K. Whitaker (eds.), *Econometric Analysis for National Economic Planning*, Butterworths, London.

Comments on: "An alternative to the mainstream model of inflation with an application to Switzerland" by Franz Ettlin

by Wilhelm Fritz

In this paper, the author takes the position that to explain the inflation process "the Phillips curve and the associated NAIRU are seriously flawed concepts". It is worth distinguishing between the NAIRU-model, which is essentially an empirical construct, and the natural rate hypothesis, which is actually the concept the author attacks. Clearly, when looking at unemployment in Switzerland since the second world war it is not obvious what the "natural" unemployment rate could be. A lengthy period of practically zero unemployment came to an end due to the oil crises and since the beginning of the 1990s, Switzerland experienced a considerable increase to now more than 5%, a level considered unthinkable by many just a few years ago.

As an alternative to a Phillips-curve specification, the author presents a price-wage model in levels and estimates it via an error-correction equation. This approach is not new. Already in the early 1960s Sargent presented a real-wage model for the United Kingdom which was extensively used until it broke down in the face of the first oil crisis. It has since been resurrected and virtually all the current macro-models in the United Kingdom have wage-price equations estimated as error-correction equations.

How does one justify a level model? As is mentioned in the paper, one argument runs along the dynamic properties of the price, wage, unemployment and GDP gap series. For a Phillips curve, both price changes and unemployment should be stationary, but they rarely are. It would be helpful if the paper contained some stationarity test.

A second justification can be derived from theory. The paper highlights the implication for the determination of the price level which are implicit in the NAIRU concept. However, instead of criticising a specific variant of the Phillips curve, the author could explain the theoretical underpinning for his own price level equation which actually seems to be a level-version of the Nordic Model of Inflation which emphasises the role of the external sector in the price formation process. Even in this context, several formulations could be discussed, all of which serve to underline the crucial role of foreign prices for smaller countries.

The author regards imports as intermediate goods and includes them in an extended production function. Alternatively – or complementing this formulation – he could have looked at the prices of export or final import goods. One could assume that the exporting firms face a price constraint which via profit maximisation naturally leads to real wages as a crucial variable. Similarly, one could assume that firms competing with final import goods face a price constraint. In the Nordic tradition, export prices have been crucial in determining domestic wage and price trends while in the French version of the Nordic model the import price channel is crucial. Anyway, it would be helpful if the model were explained. This way is would also be easier to understand why lagged import prices capture the effect of wages.

In estimating the error correction model the author uses the two step-procedure although it is now generally accepted that a one-step procedure (with the price change as the dependent variable and both first difference and level variables on the right-hand side) is more efficient. Nevertheless, the estimation results support the author's view that a level formulation outperforms the usual Phillips-curve specification and that the development of import prices has had a major influence on the Swiss inflation experience during the last two decades.

Inflation and unemployment in Belgium

Michel Dombrecht and Philippe Moës[1]

Introduction

Low inflation and low unemployment belong traditionally to the objectives of economic agents as well as of governments. It is, therefore, important to know whether both of these objectives can be attained simultaneously.

As is well known, it was originally thought that there was a negative relationship between inflation and unemployment: to reduce unemployment, the economy had to be stimulated, implying a rise in inflation. On the other hand, inflation could be brought down by increasing the number of unemployed. This negative trade-off is known as the Phillips curve. In the 1970s, most economists became convinced of the inexistence of such a trade-off, at least in the long run: the long-run Phillips curve is vertical. Only in the short run could there exist a negative correlation between inflation and unemployment, but over larger intervals of time, the latter would tend towards its equilibrium value which became known as the NAIRU (Non-Accelerating Inflation Rate of Unemployment).

These considerations have continued to inspire economists, also in the recent past, and they are of course also important for monetary policy. As far as inflation is concerned, the actual mainstream idea is to stress the importance of inflationary expectations (vertical long-run Phillips curve), whereas in the short-term, inflation may also be affected by disequilibria in the markets for products and labour (negative slope of the short-term Phillips curve). An unemployment rate larger (smaller) than the NAIRU normally implies a negative (positive) output gap; i.e. an output level smaller (larger) than potential output, which puts downward pressure on inflation. As far as unemployment is concerned, special interest developed in the NAIRU concept. If a natural rate of unemployment exists and if its determinants could be identified, then this knowledge would open a unique channel for analysing the unemployment problem and for guiding employment policy. Most of the analysis in this respect has, economically speaking, been applied to large and relatively closed economies. They are not necessarily applicable to small open economies, like the Belgian one. This paper, therefore, approaches the inflation versus unemployment question from a Belgian perspective.

1. Price formation

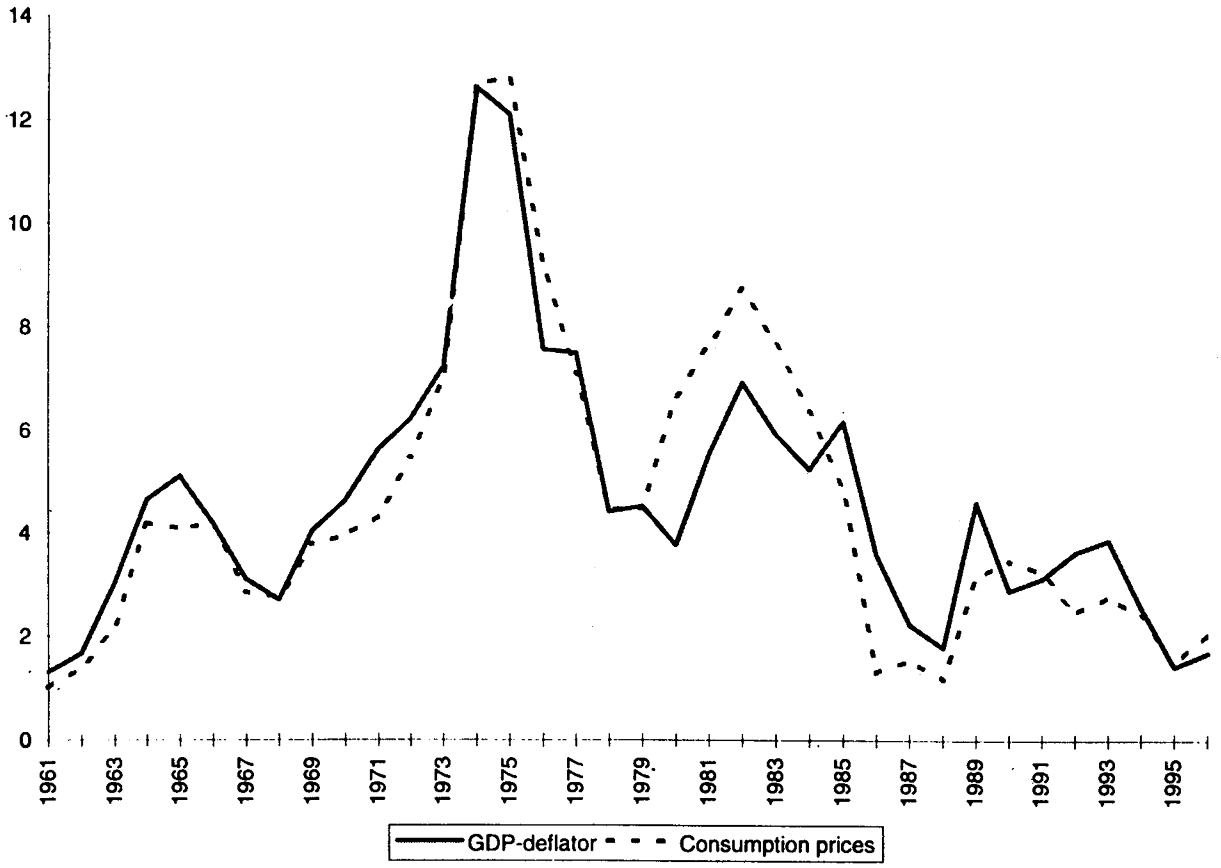
Prices can be measured at different levels and following different methodologies. The consumer price index, which represents a fixed basket of goods and services and is quickly available, is, besides adjustment for indirect taxes, the result of a weighted average of prices of domestically produced goods and services set by domestic producers and of prices of imported goods and services set by foreign producers and converted into domestic currency. The prices set by domestic producers are themselves output prices which are determined by the costs of labour, capital, intermediate inputs (a large part of them being imported) and by a profit margin.

---

[1] The authors are members of the Research Department of the National Bank of Belgium. The views expressed in this paper are their own and do not necessarily reflect those of the Bank.

The gross domestic product (GDP) deflator is a national accounts concept which is germane to an output price concept except that it not only relates to private but also to public output and that the direct influence of intermediate input prices has been removed. The GDP deflator constitutes an important link in the explanation of consumption prices since it reflects the domestic origins of inflation (wage costs, capital costs and the profit margin). This paper focuses on the GDP deflator. Notwithstanding the conceptual differences, GDP inflation is strongly correlated with consumer price inflation as is shown in Chart 1.

Chart 1
Inflation in Belgium
Percentage changes



1.1 Historical overview

The inflation rate in Belgium has followed an upward trend since the beginning of the 1960s up until 1974 – the outbreak of the first oil price shock (Chart 2). Since then, inflation has followed a downward trend, which continued in the 1990s. In 1995, inflation was at a level equal to the historical low levels of the early 1960s. A similar evolution can be found in the OECD area and in Belgium's three neighbouring countries, which are also our main trading partners. In the 1990s a very strong convergence of inflation rates occurred in Belgium and its neighbours (Chart 3).

Chart 2

Inflation in Belgium and in the OECD

GDP deflators, percentage changes

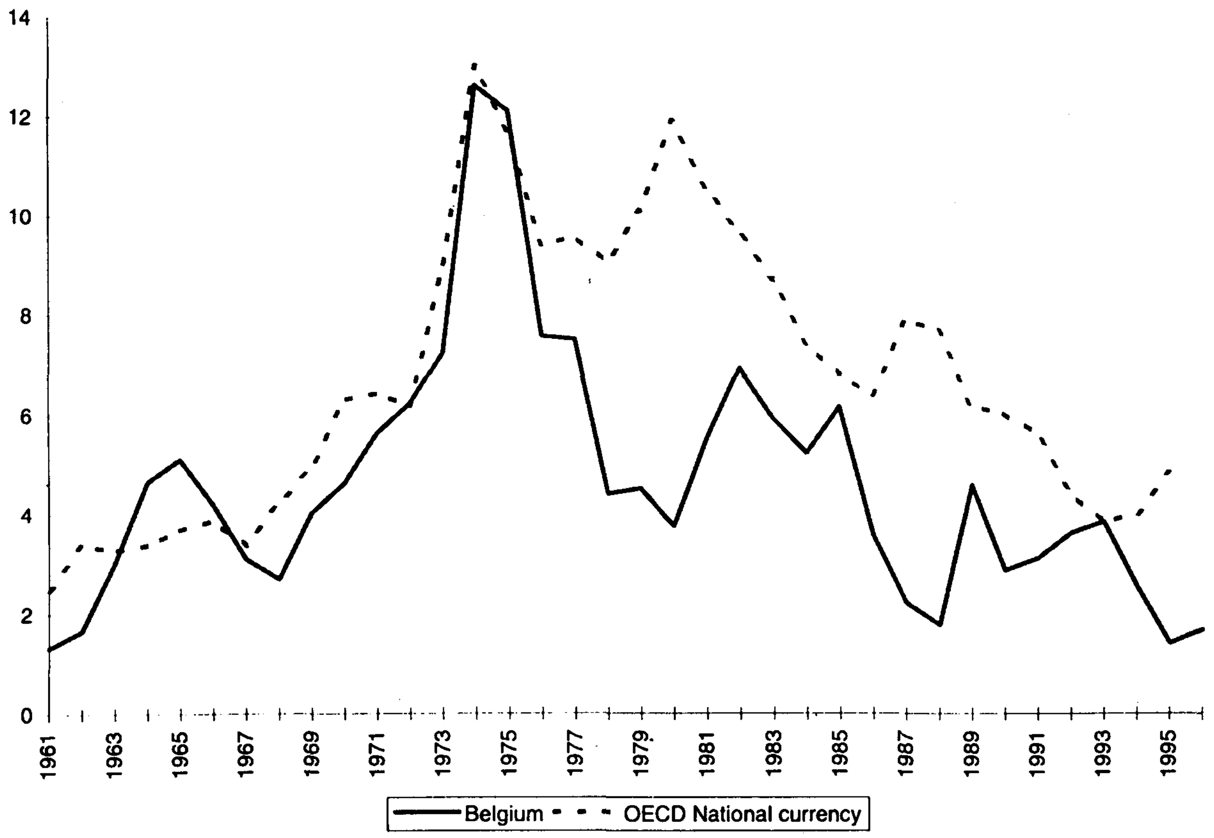


Chart 3

Inflation in Belgium and neighbouring countries

GDP deflators, percentage changes

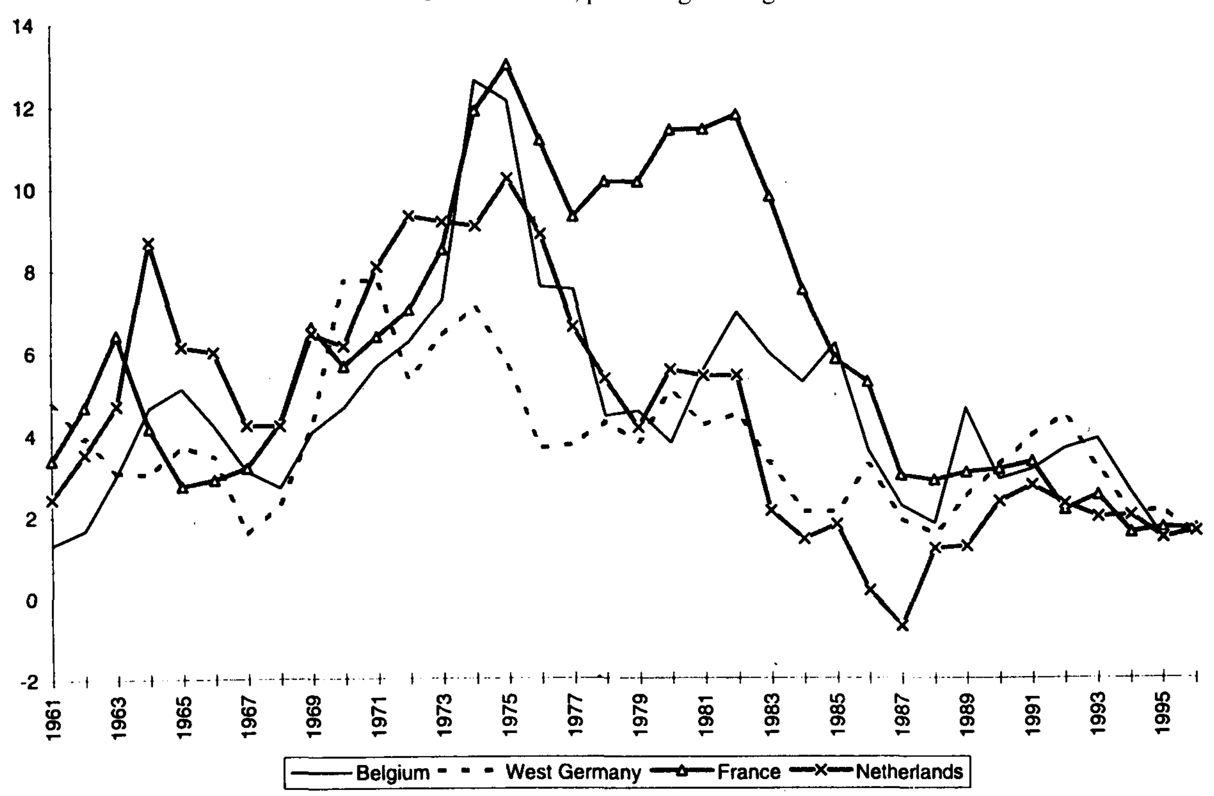


Chart 4 shows the year-on-year change of the inflation rate. It demonstrates that inflation volatility is not constant. This is confirmed in Table 1, containing averages and standard deviations of inflation in each decennium. The table shows a strong negative correlation between the level of average inflation and its volatility. For the monetary authorities, this is an important argument to focus on for the following reasons:

- high inflation entails large price volatility, implying a high degree of uncertainty for producers. Investment, process and product innovation and employment fare better when producers can more easily predict the future cash flows to be expected from such new projects;
- stable prices render relative price changes between different goods and services more visible to consumers and producers. Because high inflation is correlated with high volatility and uncertainty, market signals are less clear and, therefore, the efficiency of the market mechanism is blurred, which may lead to welfare losses;
- questionnaires[2] undertaken in different countries show a strong inflation aversion of the population. The main reason for people's concern about inflation is the perception that it hurts their standard of living. This may be due to the non-neutral character of taxation with respect to inflation. In most countries tax brackets are not being indexed to price changes and inflation accounting is not a part of standard accounting principles.

Chart 4
Year-on-year change of inflation in Belgium
In percentages

---

2 See, for example, Shiller, R. J. (1996): "Why Do People Dislike Inflation?" NBER, *Working Paper* No. 5539.

Table 1
Inflation: averages and volatility

Period	1961-1996		1961-1969		1970-1979		1980-1989		1990-1996	
	AVG	STD	AVG	STD	AVG	STD	AVG	STD	AVG	STD
Belgium	4.63	2.58	3.31	1.30	7.23	2.95	4.56	1.70	2.71	0.92
Western Germany	3.76	1.65	3.33	0.97	5.52	1.64	3.0	1.20	2.89	1.06
France	6.07	3.55	4.25	1.43	9.29	2.42	7.15	3.69	2.26	0.71
Netherlands	4.46	2.95	5.14	1.86	7.66	2.01	2.33	2.27	2.03	0.44
OECD	6.70	2.84	3.62	0.68	9.05	2.28	8.30	1.88	4.76	0.88

AVG = average; STD = standard deviation.

1.2 Long-run determinants of inflation

Section 2 of Annex 1 describes a model of optimal price setting in a market characterised by monopolistic competition. In such a market constellation producers take into account the degree of price elasticity of demand for their products when fixing prices. If the price elasticity of demand is small, firms can charge high mark-ups above production costs and vice versa. In the limit (perfect competition), the price elasticity is infinite and mark-ups disappear.

Consumers decide on their optimal consumption-saving behaviour as well as on the optimal mix of their consumption basket. When relative prices change, consumers will reallocate their consumption portfolio and the extent of this restructuring depends on the price elasticities of demand for all products. These price elasticities depend on:

- the clearness of market signals, mentioned above, which increases with the degree of overall price stability;
- the market structure; a cartel market limits the consumer's freedom of choice and hinders the composition of an optimal consumption basket;
- the degree of openness of the economy which determines the accessibility of the consumer to the international markets of goods and services. If domestic producers are largely exposed to competition with foreign firms both on domestic and foreign markets, their mark-ups will tend to be smaller. Contrary to the case of large and relatively closed economies, this consideration should be taken into account in the price formation process of firms in a small economy such as Belgium.

On the basis of these arguments and assuming that consumers allocate their consumption according to an Almost Ideal Demand System (AIDS), Annex 1 explains the GDP deflator by:

- unit labour costs (nominal wage cost per unit of output), which, in turn, depend on wage cost per employee (further explained in Section 2) and apparent labour productivity;
- a mark-up above unit labour costs, the extent of which depends on the price elasticity of demand. At the macro level, the price elasticity for the country as a whole, is related to the market share of domestic producers and hence to the relative price of foreign with respect to domestic firms converted into a common currency. This relative price is nothing else than the real effective exchange rate. When foreign prices are relatively high, market shares of domestic firms are relatively high as well, such that they can raise their own prices in order to reach the optimal profit maximising mix between profit margin and market share. This analysis demonstrates that even the GDP deflator, which reflects the domestic origins of inflation, is

subjected to a *direct* influence of foreign prices. This result is, of course, especially relevant in the case of countries where foreign trade represents a large proportion of final demand and is absent in the inflation analysis originating in large countries. It should be added that foreign prices also exert an *indirect* influence on domestic prices. Prices of imported consumer goods are part of the overall consumer price index, a variable taken into account in wage negotiations. Furthermore, imported intermediate inputs are part of production costs and an integral part of output prices of domestically produced consumer goods. These indirect effects do not play their role through the profit margin, but via nominal wage costs.

- The analysis of the inflation process, therefore, demands insight into the course of both nominal unit wage costs and the profit margin. If the latter cannot be sufficiently explained, it will be hard to understand and interpret the inflation process (not only on the basis of the GDP deflator but also other price indexes). Monetary policy in many large economies exerts its influence through its effects on inflationary expectations of economic agents. More specifically, credible monetary policy may contribute to wage moderation. In Belgium, wages are indexed to consumer prices such that expectations play only a minor role. Furthermore, in such a small open economy, foreign prices and exchange rates are principal factors underlying domestic inflation. Chart 5 compares unit labour costs with the GDP deflator in Belgium. It is clear that the GDP deflator can diverge for rather long periods from unit labour costs, thereby giving rise to marked changes in the macroeconomic profit margin as shown in Chart 6. According to the theoretical analysis, the profit margin is related to relative prices or, at the macro level, the real effective exchange rate of the Belgian franc. An increase (decrease) in the relative price should be interpreted as a real effective depreciation (appreciation). Although somewhat blurred during 1982-86 by the strong dollar and oil price movements, Chart 7 demonstrates the positive relationship between the mark-up and the real effective exchange rate, based on GDP deflators. This implies that foreign prices constitute a constraint on price setting by domestic firms. The pass-through from domestic unit wage costs to output prices is incomplete: if domestic wage costs increase, while foreign prices remain constant, only part of the additional labour costs can be transferred to price increases and the profit margin will decline.

Summarising, domestic output prices are determined as a function of a weighted average of the foreign price level, expressed in domestic currency, and unit labour costs (Chart 8), with the latter consisting of two components: wage costs per employee (W), which will be analysed in Section 2; the inverse of apparent labour productivity (L/Y) which itself can be decomposed into: influence induced by substitution between factors of production (due to changes in relative factor prices): (L/K) and Total Factor Productivity (TFP), i.e. the autonomous technological progress.[3] The employment-capital ratio (L/K) can itself be further decomposed, as employment (L) is the product of the active population (L^S), reduced by the proportion out of work (U): $L = L^S(1-U)$.

The econometric analysis in Annex 2 evaluates the weights of all these factors that influence inflation, and derives the following long-term steady-state inflation equation:

$$\overset{\circ}{P} = 0.62\overset{\circ}{W} + 0.44\overset{\circ}{P}{}^{*} - 0.06\overset{\circ}{OIL} - 0.65\overset{\circ}{TFP} + 0.19\left(\overset{\circ}{L^S / K}\right) - 0.002\Delta U$$

where: $^\circ$ = growth rate of the relevant variable;

P = GDP deflator;

W = nominal wage cost;

P^* = weighted average foreign GDP deflators expressed in the same currency as P;

---

3 In the case of a Cobb-Douglas production function: ($Y = (TFP)L^{\alpha}K^{(1-\alpha)}$), the employment intensity of output (i.e. the inverse of labour productivity) can be written as: $L/Y = (1/TFP)(L/K)^{1-\alpha}$.

Chart 5
GDP labour and wage cost per unit of output in Belgium
Index numbers, 1960=100

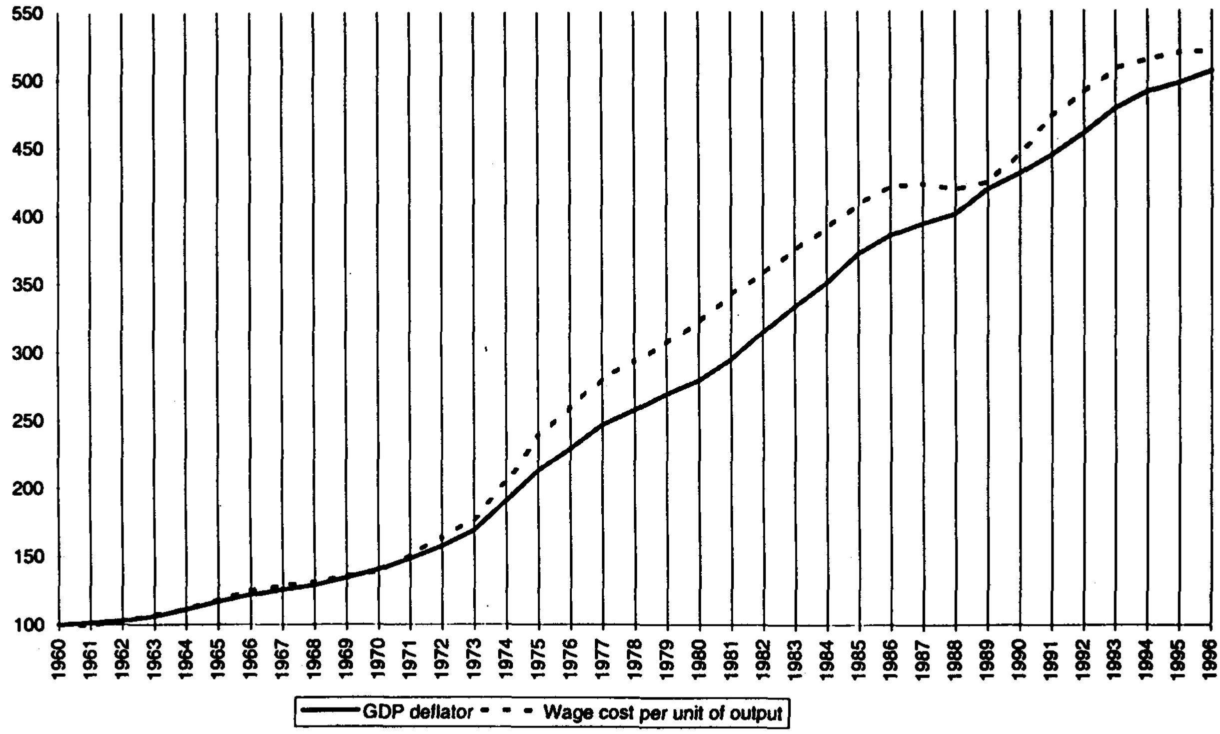


Chart 6
Macroeconomic profit margin in Belgium
Index numbers, 1960=100

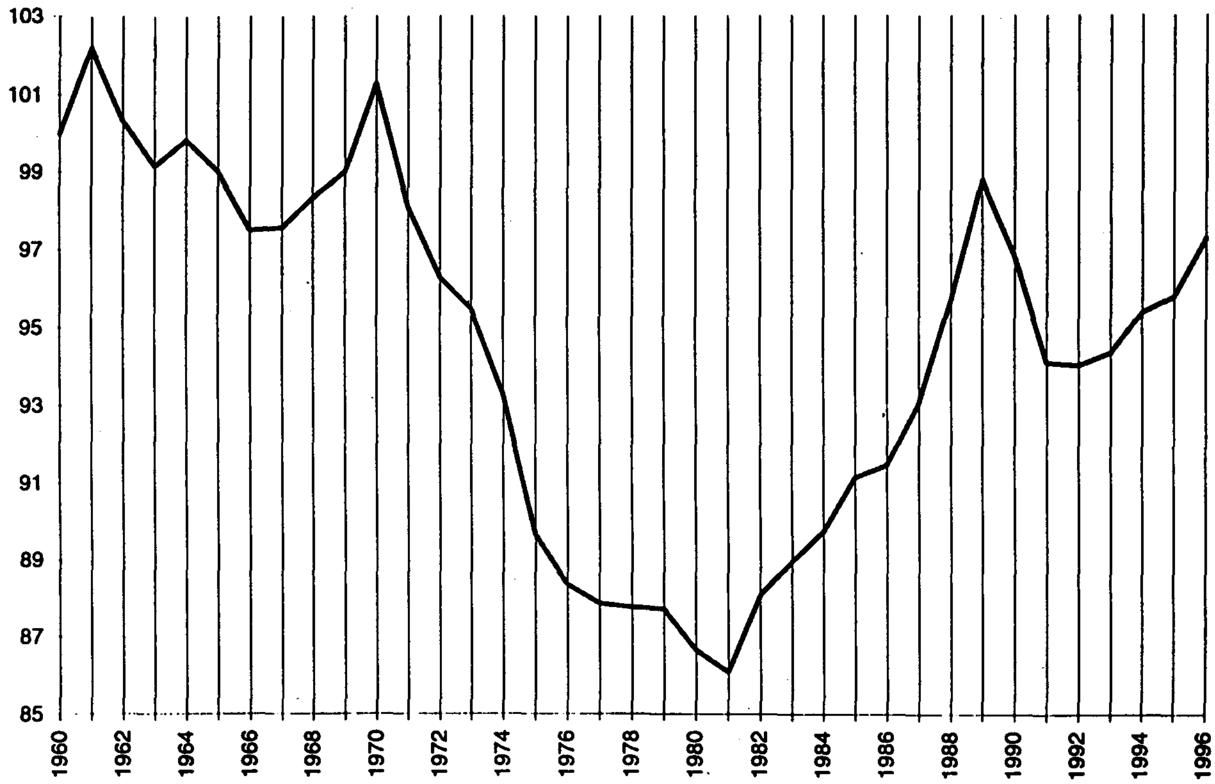


Chart 7
Profit margin and relative GDP deflator in Belgium
In logarithms

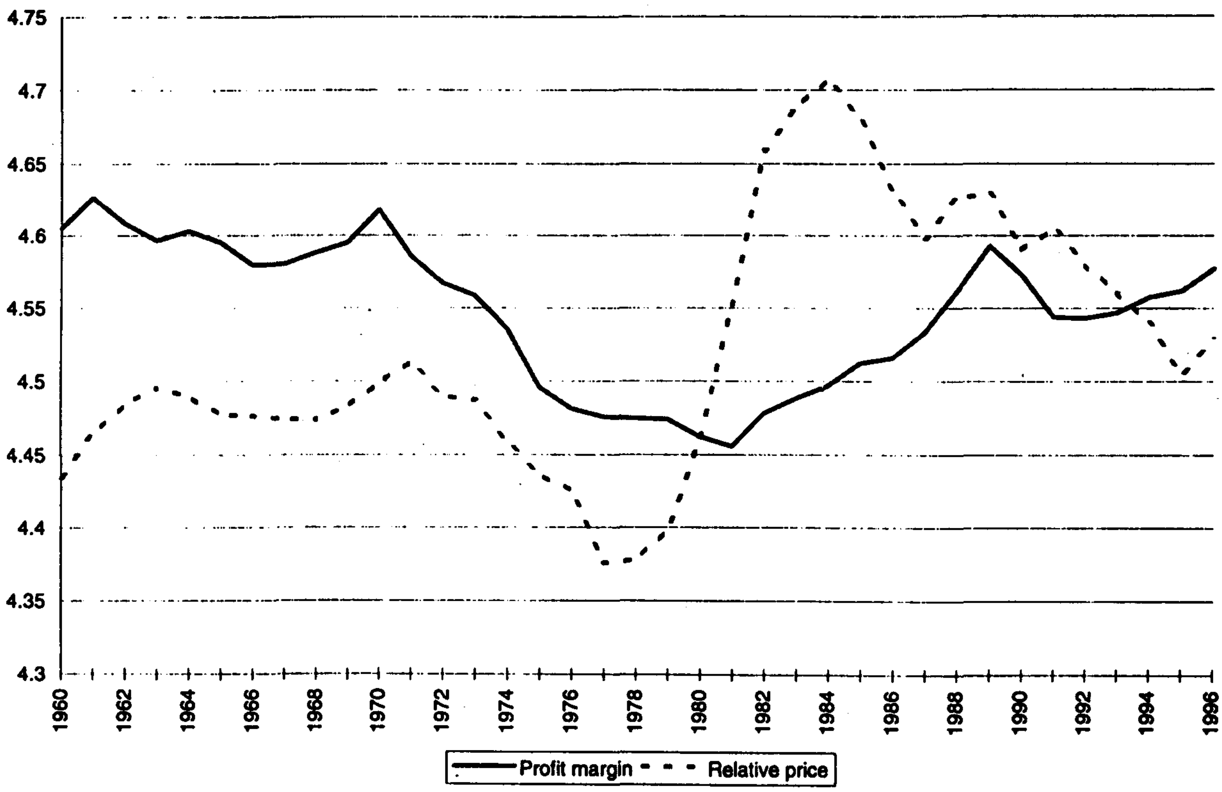
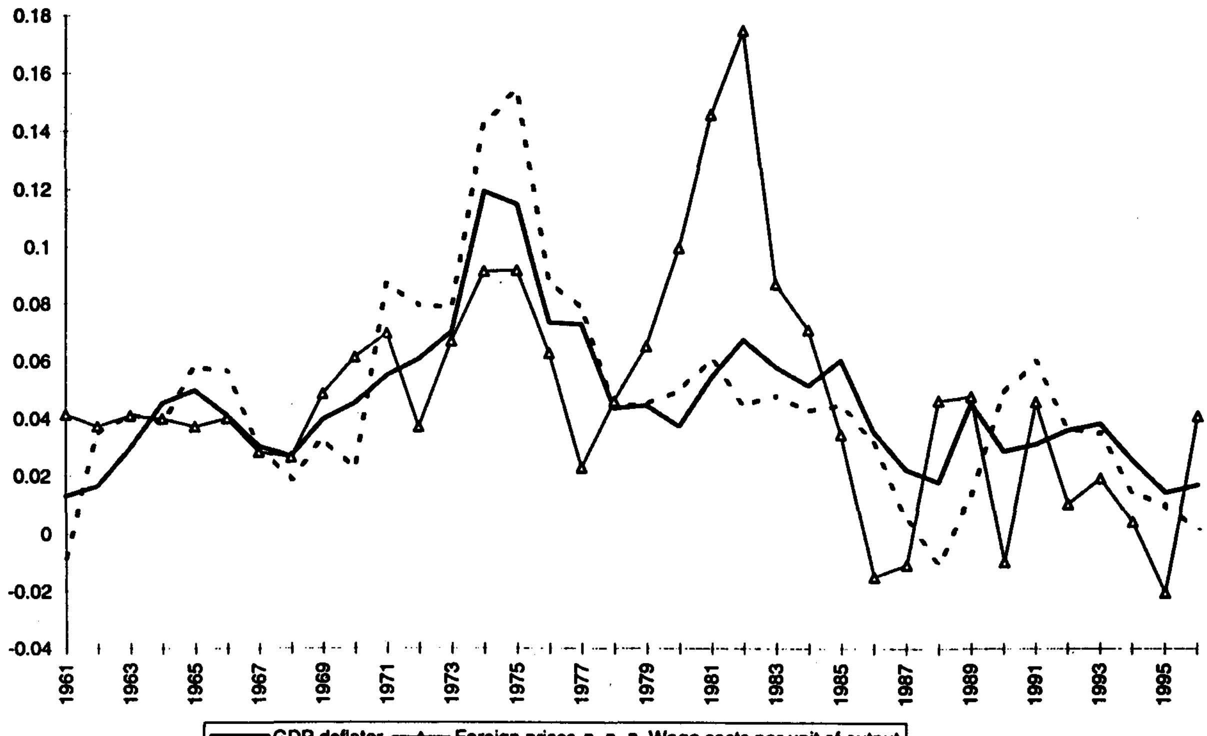


Chart 8
Explanatory variables of GDP inflation
In annual percentage changes



OIL = index of world oil prices;

TFP = total factor productivity;

L^S = active population;

K = capital stock;

U = unemployment rate.

The results indicate that the direct weight of foreign prices in domestic inflation is quite substantial. Oil price increases somewhat moderate this foreign influence as foreign countries that are producers of oil do not affect the Belgian market share because its share in those products is zero anyway. The part of foreign inflation that is connected with oil price movements is thus irrelevant for the domestic inflation rate. We may, therefore, conclude that foreign inflation has a direct weight equal to 38% in domestic GDP inflation, as compared to 62% for unit labour costs. By implication, foreign prices may drive an extensive wedge between domestic prices and unit labour costs.

1.3 Short-run determinants of inflation

In the short run, inflation reacts with a certain delay to its long run determinants. In the framework of a prediction exercise, it is important to take these reaction lags into account, while they are far less relevant in a longer-run analysis of inflation. Annex 2 contains the estimated short-term inflation equation.

2. Wage formation

Except in periods of government intervention, the formation of wages is the result of negotiations between employers and unions, as explained in Section 1 of Annex 1. The unions' objective is to obtain the highest possible after-tax net real wage; i.e. the real wage cost after correction for the tax wedge (expressed as the difference between wage cost for the firm and net wage income earned by the employee in percentage of wage cost). During the negotiations, the unions anticipate the reaction of employers to their wage demands. They take into account the probability that negotiations fail or are followed by lay-offs or that those actually unemployed may fail to find a job. This probability is proportional to the observed unemployment rate. Unions also are aware that when becoming unemployed members will earn unemployment benefits, the magnitude of which, therefore, affects their wage demands. After conclusion of the wage negotiation, firms decide on their (profit-maximising) prices, output and production technology; i.e. the optimal mix of factors of production. The result of the wage negotiation will, therefore, affect employment.

Taking account of these considerations, wage formation depends on:

- unemployment benefits: they constitute a safety net in case of becoming unemployed if wages turn out to be too high to preserve a high level of employment. The more generous this safety net, the lower the perceived opportunity cost of loosing one's job and, therefore, the higher wage demands will tend to be;
- the tax wedge: an increase in employers' social security contributions raises wage cost, and a rise in taxation on employees will induce the latter to raise their wage claims;
- the market share (and hence relative prices) of domestic producers on world markets: higher market share tends to be accompanied by higher profitability and stimulates the quest for higher wages;
- the unemployment rate: the higher this rate, the higher the perceived probability of becoming unemployed, which will exert a moderating influence on wage claims.

From the econometric estimation results, the following long-term wage growth equation can be derived:

$$\overset{\circ}{W} = \overset{\circ}{P} + 0.31\left(\overset{\circ}{P^*} / P\right) + 0.63b - 1.22\left(1 - \overset{\circ}{\tau}\right) - 0.026\Delta U$$

where b = real unemployment benefit per person unemployed;

τ = tax wedge;

ΔU = change in the rate of unemployment, expressed as a percentage.

This long-term equation shows a negative dependence of wage growth on changes in the unemployment rate; an increase of the unemployment rate by 1%, ceteris paribus, reduces wage growth by 2.6 percentage points. This coefficient is generally regarded as a measure of wage flexibility since it reflects the magnitude of price reaction to disequilibria between demand and supply, in casu of labour. A few remarks should be made in this respect. First, in Belgium, this flexibility was frequently enforced by government intervention in the wage formation process.[4] Second, despite the observed flexibility of wages to changes in the unemployment rate, the growth of wage costs has been continuously stimulated by the ever increasing tax wedge. Third, the significant rise in real unemployment benefits in the 1970s has contributed to the strong growth of wages in that period.

3. The ultimate determinants of domestic inflation

As the preceding analysis has shown, steady-state inflation and wage growth are interrelated. Taking account of all interactions (i.e. calculating the reduced form of the inflation equation) results in the following expression of the ultimate determinants of inflation:

$$\overset{\circ}{P} = 1.10\overset{\circ}{P}{}^* - 0.10\overset{\circ}{POIL} + 0.67b - 1.32(1 - \tau) - 0.032\Delta U - 1.13\overset{\circ}{TFP} + 0.34\left(\overset{\circ}{L^S} / K\right)$$

The ultimate factors explaining inflation are, therefore:

- the foreign rate of inflation: if foreign inflation rises by 1 percentage point, domestic prices also accelerate by 1 point. This implies, of course, a strong dependence of domestic prices on exchange rates, especially with respect to Germany, France and the Netherlands, our main trading partners;
- total factor productivity: an increase in *TFP* raises labour productivity, which reduces unit wage costs and puts downward pressure on prices;
- the per capita capital stock: an increase in this variable, with a constant unemployment rate, also implies a reduction in unit labour costs and therefore reduces inflation;
- the tax wedge: when rising, this induces an upward movement of wage costs (both directly if employers' social security contributions go up and indirectly through the negotiated gross wage rate) and hence on prices;
- real unemployment benefits: an increase in this variable reduces the opportunity cost of becoming unemployed and encourages higher wage demands;

---

4 These interventions took the form of real wage freezes, imposition of wage norms, limiting the indexation mechanism, changing the reference price indicator to which wages are indexed. This is no longer the general consumption price index but the so-called health index.

- the unemployment rate: this determines the perceived probability of losing one's job. It puts downward pressure on negotiated wages and on prices.

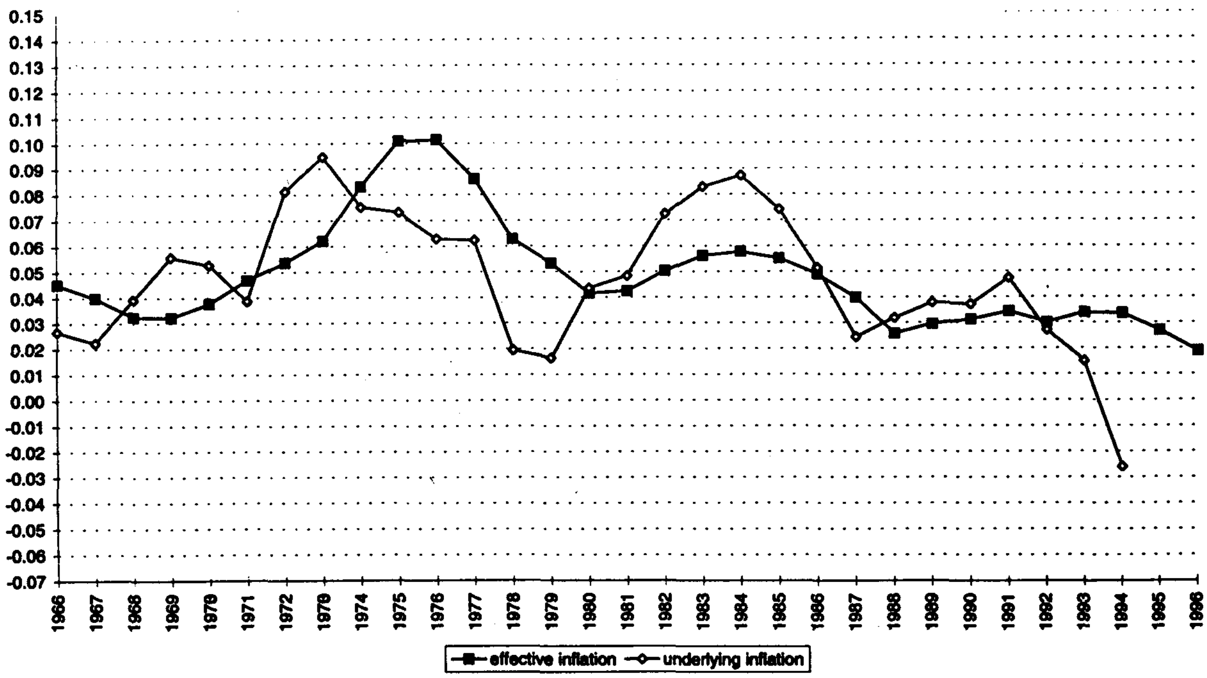
The previous equation allow us to compute an underlying inflation rate based on the growth rates of the explanatory variables. Substituting historically observed values for these variables allows us to compute the underlying inflation for each year. However, important shocks to the exogenous factors immediately and strongly affect the underlying rate of inflation whereas observed inflation adapts only slowly. Hence the use of moving averages over three years in Chart 9 where the effective and underlying inflation rates are shown. It appears that the underlying inflation rate is a leading indicator of inflation. Although the former seems to be too volatile, the turning points are well explained, showing a lead of about two years.

What have been main factors behind underlying inflation in the last thirty years? To make things simple, we aggregate the many variables given by the theoretical analysis and decompose progressively. Chart 10 gives the foreign inflation rate and the impact of all the remaining factors taken together. Until the mid-1970s the trend in Belgian inflation reflects the foreign one. The other factors are erratic around zero. But between 1978 and 1984, their contribution exceeds -2%. and even more between 1981 and 1983. As a consequence, the underlying inflation falls below imported inflation. In the second half of the 1980s, foreign inflation and the other factors have more or less the same positive impact, between 1 and 3%. From 1992 onwards, the other factors fall sharply together with the underlying inflation which in 1994 is below the foreign inflation rate. To conclude on this matter, even if the domestic inflation is indexed to the foreign one, the domestic factors are able to introduce quite significant and persistent divergencies. We now analyse more closely the contribution of these other factors.

Chart 9

Effective and underlying inflation rates

Moving averages over three years



We can distinguish between one external variable (*OIL/p**) and internal ones: real unemployment benefits, the tax wedge, *TFP*, capital stock per capita and the unemployment rate. We further distinguish between permanent and transitory internal factors. Permanent factors are those whose growth rate may permanently lie above zero: unemployment benefits, *TFP* and capital stock per capita; in contrast the tax wedge and the unemployment rate should stabilise somewhere in the long run. Hence, their contribution can only be transitory. Chart 11 gives the contributions to the inflation differential according to this subdivision. The permanent factor is the most important one over the whole period. The transitory factor is less important until 1983 and the differential follows the "permanent" fluctuations plus some impact from the oil price. From 1984 onwards, the impact of the transitory and permanent components are comparable but the volatility of the permanent factor goes down and the differential follows the "transitory" fluctuations as well as the impact of the reverse oil price shock.

Chart 10
Underlying and foreign inflation rates
Moving averages over three years

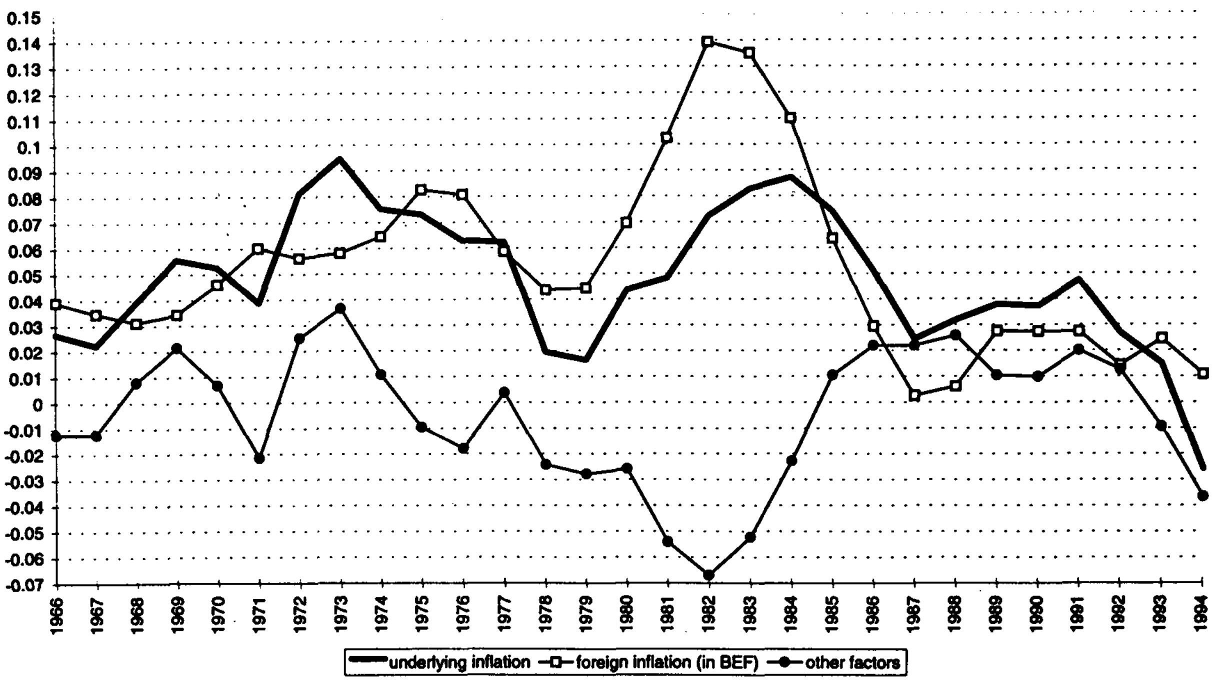
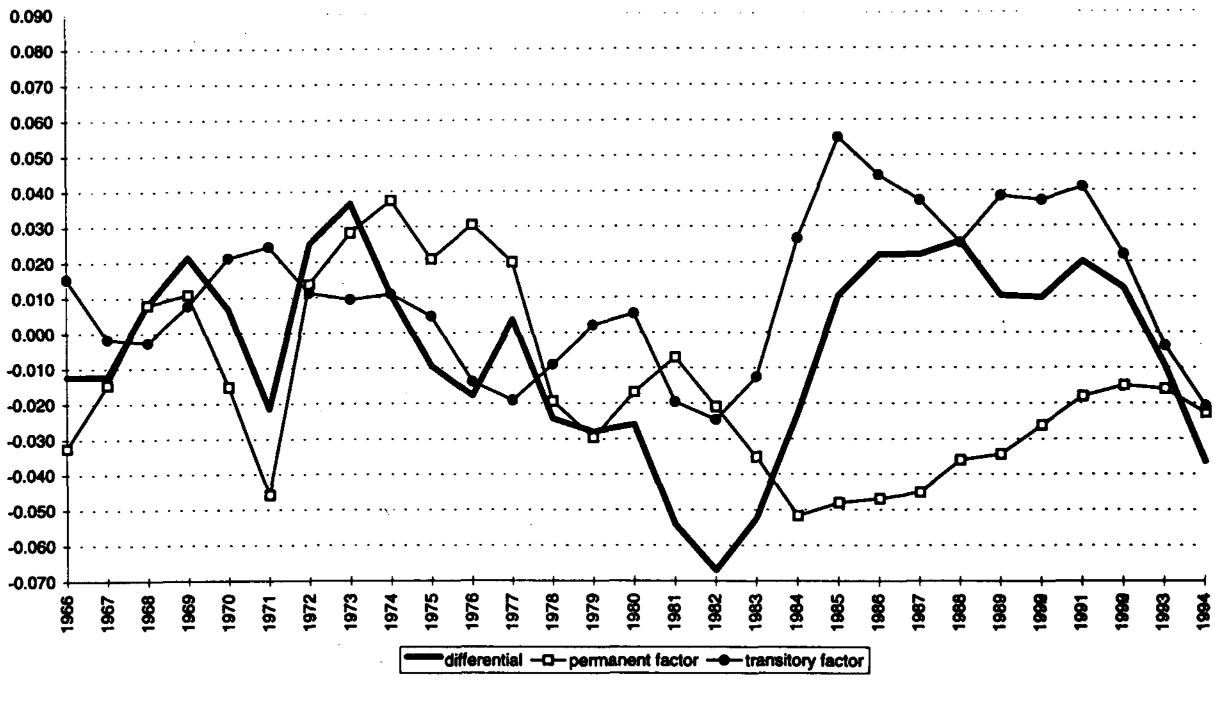


Chart 11
Infaltion differential Belgium – foreign competitors
Moving averages over three years

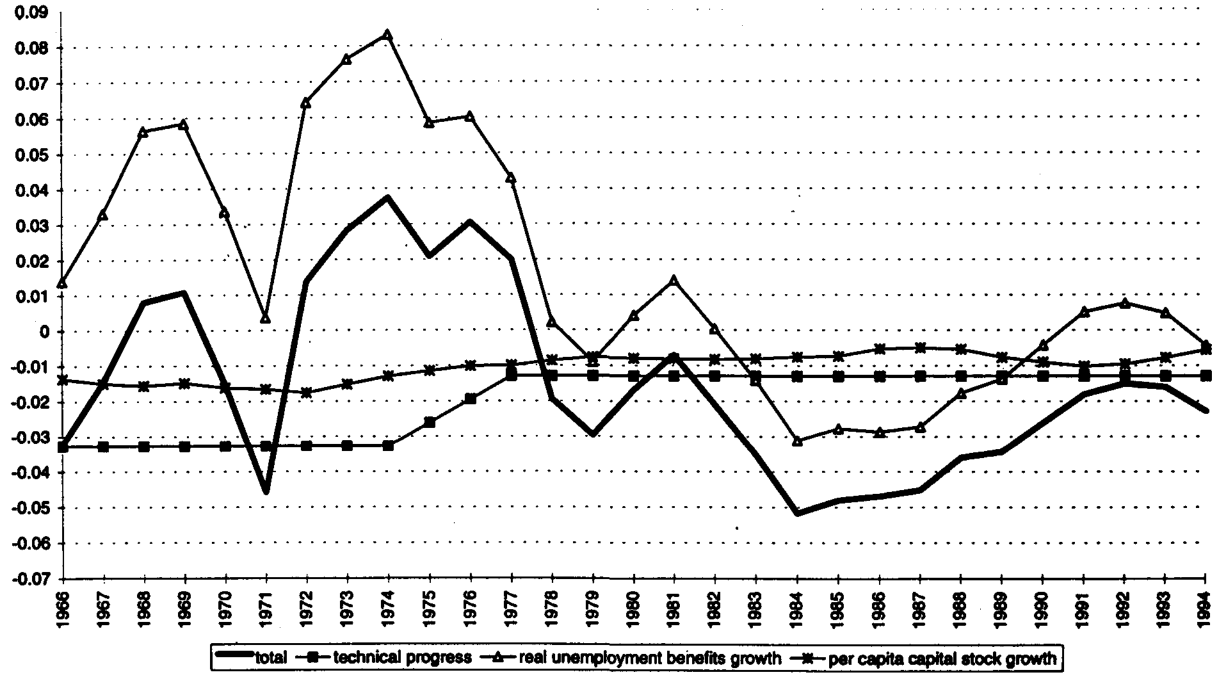


Behind the permanent factor (Chart 12) is the growth of unemployment benefits adjusted for productivity growth (*TFP* plus capital per head). In fact, the estimates (0.67 for the unemployment benefits and 1.13 for productivity) imply an overcorrection for productivity. Until 1971, the benefits had no link to previous wages but depended on age and gender. Real growth adjustments were only occasional but quite significant (see Chart 13). They happened in 1967, 1968 and at the junction to the present system, in 1972. These shocks explain the two first rises in the unemployment benefits contribution. Afterwards, a single shock appeared in 1975, probably as a consequence of the massive arrival of new and high-wage unemployed persons. This shock added to the previous 1972 real adjustment. In the opposite direction, a fall in real unemployment benefits occurred in the middle of the 1980s. These evolutions underlie the movements in the permanent factor. The productivity variables are very stable. The only event to mention is the 1975 productivity slowdown. *TFP* fell from 2.9% a year to 1.2%, lowering the negative contribution to the inflation differential from 3.3% to 1.3%.

Chart 12

Factors behing the permanent factor

Moving averages over three years



The transitory factor (Chart 14) follows the unemployment rate movements with one single exception: the rapid rise of the tax wedge in the early 1980s. The contribution of the unemployment rate may exceed 3% in absolute value. The tax wedge contribution is fairly constant, at +2%.

It appears that the underlying Belgian inflation rate can substantially diverge from inflation abroad. Two main factors have, in the past, contributed to discrepancies: the growth of real unemployment benefits and unemployment rate movements. However, structural shocks affected the growth of unemployment benefits in the first half of the sample and they should not repeat themselves. In the long run the benefits should follow productivity. In that case, their impact would be much less important. As far as unemployment is concerned, its fluctuations should remain transitory since a continuous divergence always calls for counteracting measures.

Chart 13
Unemployment benefits
1991=100, in logarithms

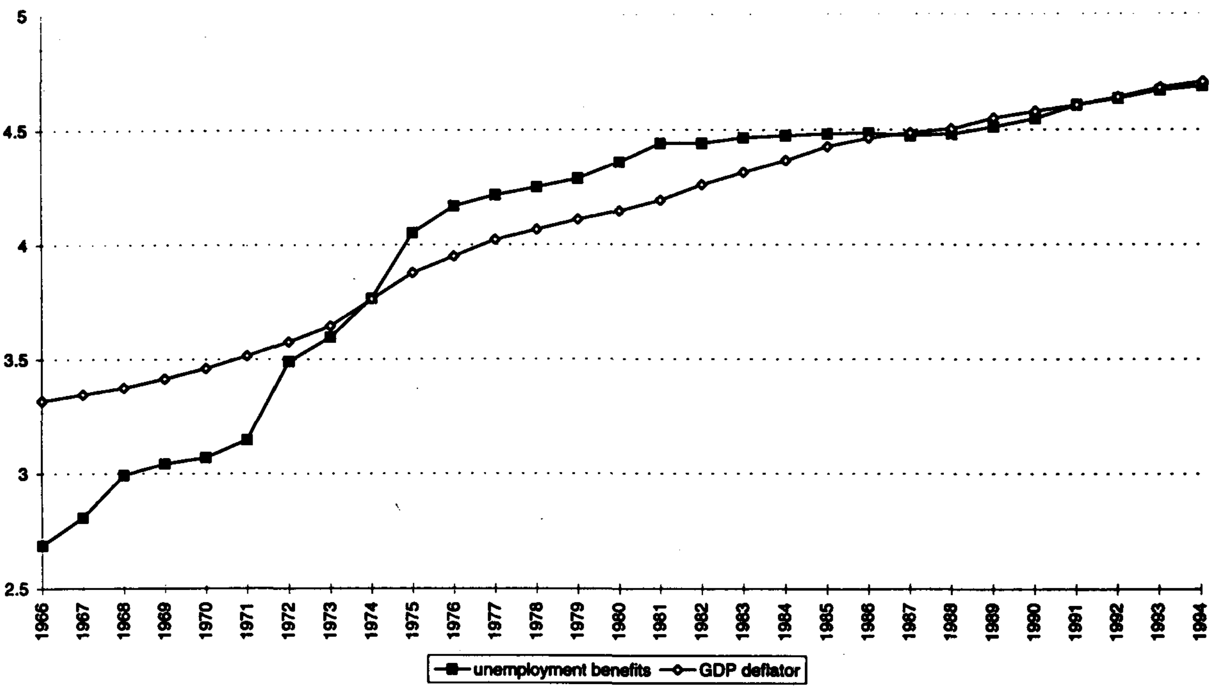
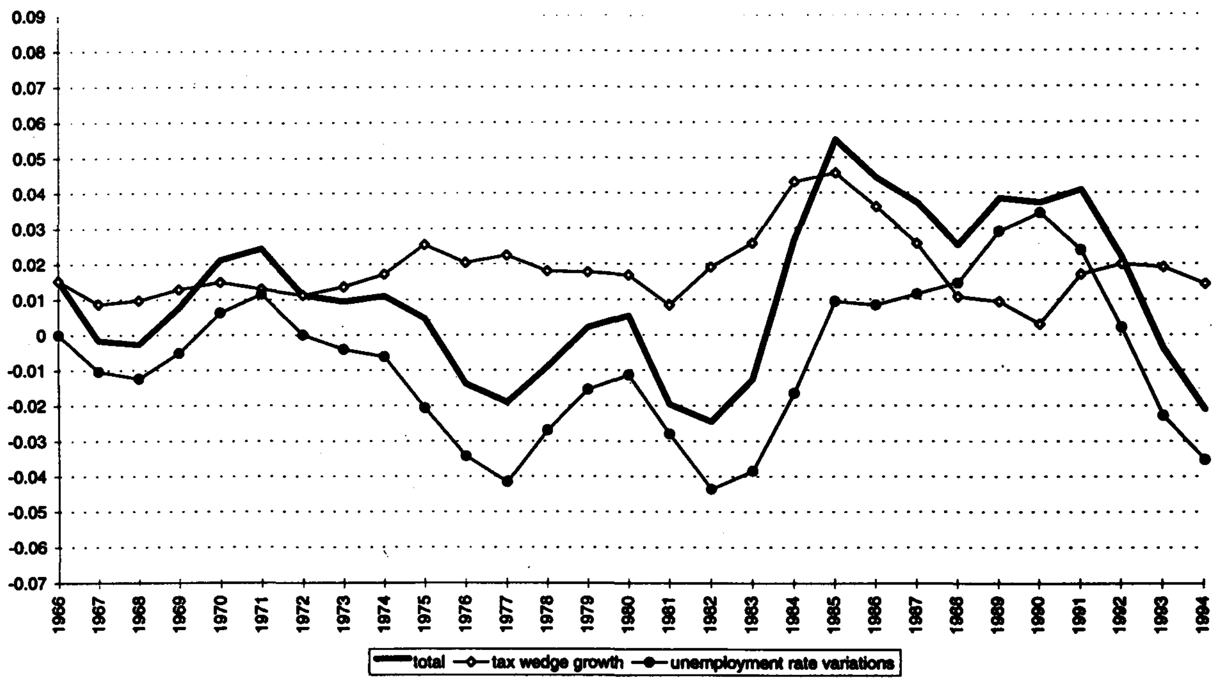


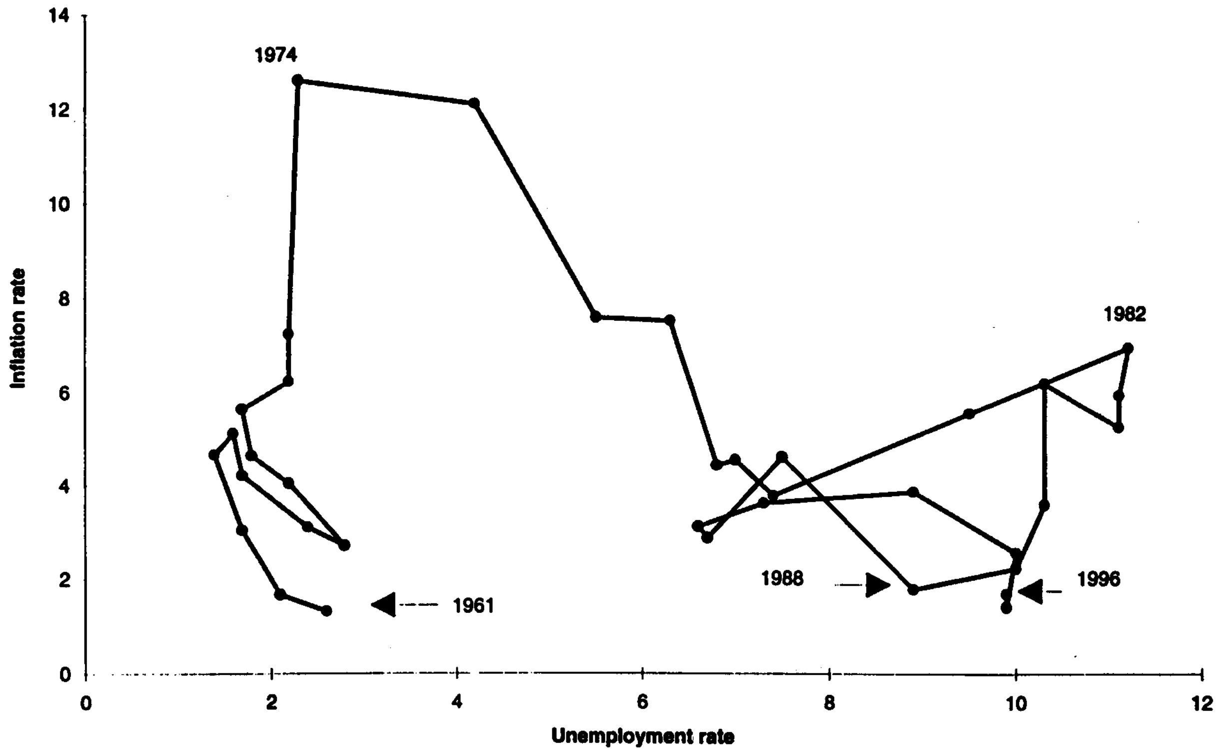
Chart 14
Factors behing the transitory factor
Moving averages over three years



4. The rate of unemployment

The reduced-form inflation equation mentioned above implies a long-run relationship between the deviations of domestic with respect to foreign inflation rates and variations of the unemployment rate. The existence of a NAIRU, on the other hand, requires long run independence of inflation with respect to both the level and variation of unemployment. In this sense no NAIRU for Belgium can be derived. Chart 15 plots the relationship between inflation and unemployment and demonstrates that the unemployment rate does not fluctuate around a relatively constant equilibrium value: the unemployment rate fluctuated around some 2% up till 1974 but shifted from 1975 onwards to about 11% in 1982. Since then, it has fluctuated within a very broad interval around some 9%.

Chart 15
Inflation-unemployment trade-off in Belgium

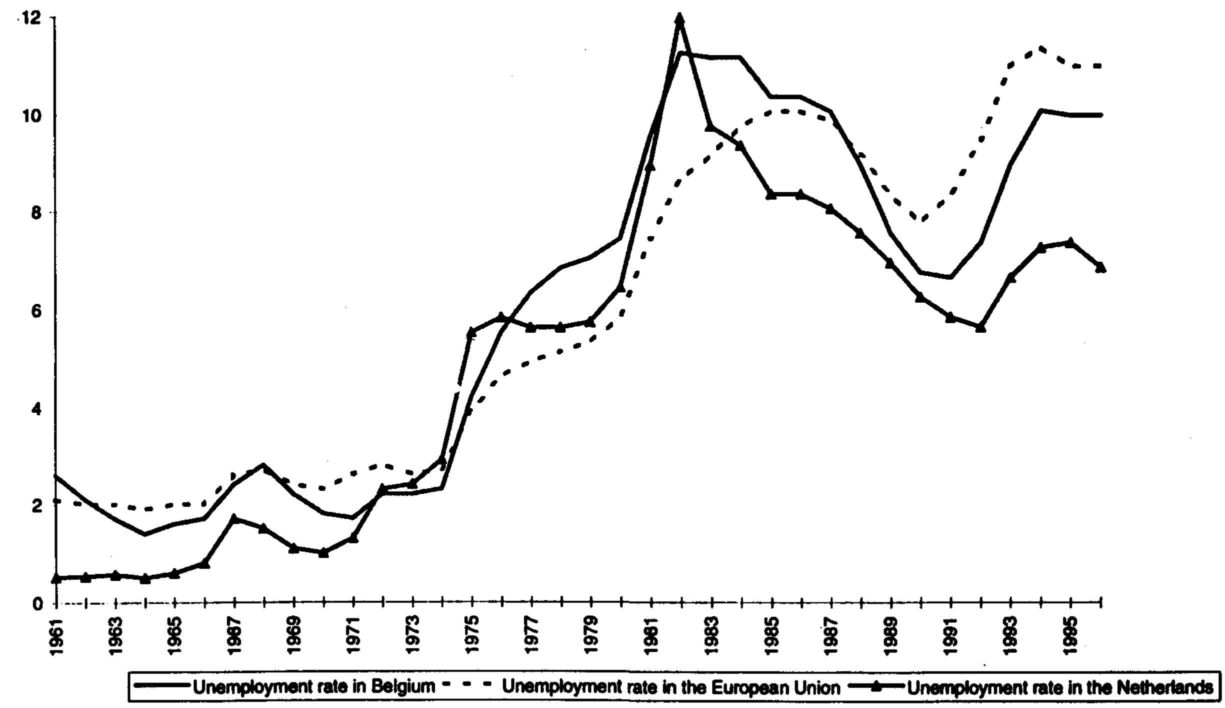


In the absence of a NAIRU, the above-mentioned price equation is not sufficient to detect the explanatory factors behind the marked shift in the unemployment rate. For that purpose, the inflation equation, which represents the supply side of the economy, has to be supplemented with a description of the demand side. This leads to the recognition that the domestic unemployment rate is not independent of foreign unemployment and, more specifically, of the unemployment situation in the rest of the European Union. Annex 3 demonstrates that in the long run the Belgian unemployment rate is determined by:

- the foreign unemployment rate. Chart 16 shows the comovement of unemployment rates in Belgium and in the rest of the Union.
- besides the foreign unemployment rate, domestic factors may play a significant role as the Dutch example so clearly demonstrates. The more extensive use of part-time employment may in part explain the lower unemployment rate in the Netherlands, although other domestic elements may also be part of the explanation. According to our analysis, these other domestic factors are related to the wage-price formation processes:

- productivity developments tend to limit the pressure on prices such that, external conditions remaining equal, demand for domestic products rises, thereby encouraging domestic employment. On the other hand, rising real unemployment benefits put upward pressure on wage claims and on inflation, thereby depressing employment. Our calculations imply that the overall contribution of these three variables on average had a positive influence on domestic employment during the period 1962 to 1994. This beneficial influence was, however, more than compensated by a fourth domestic factor:
- the development of the tax wedge, which has put upward pressure on wage costs.

Chart 16
Unemployment rates in Belgium, the EU and in the Netherlands



The analysis can be interpreted as follows. Increasing involuntary unemployment implies a growing disequilibrium in the labour market, which was not eliminated because wage formation was not sufficiently flexible. The question, however, is: why has wage formation been that inflexible? In the public debate frequent reference has been made to institutional factors such as minimum wages, inflexible working hours, wage indexation, organisation of wage negotiation, etc. Our analysis implies that the ever growing tax wedge has contributed to the explanation. Because of the lack of spontaneous and sufficient response of wages to the unemployment rate, the government had to intervene frequently in the wage formation process. At the end of periods of imposed wage moderation, wages tended to recover previously lost grounds, such that government had to intervene again, all of this giving rise to a stop-go process in wage formation. Modulating the tax wedge is probably an effective instrument to enhance the flexibility of wages towards disequilibria in the labour market. This avenue, which became part of public policies (in the form of several new measures to reduce employers' social security contributions), is, of course, constrained by the need for fiscal consolidation. This demonstrates again that a sound budget balance is a necessary condition for fiscal policy to be an effective instrument of economic policy.

Conclusions

The GDP-deflator, which is strongly correlated with the consumer price index, is in the long run determined by unit labour costs and a profit margin. The profit margin is not constant (not even in the long run) since it depends on the price elasticity of demand, which itself is related to the market share of domestic producers and therefore to the relative output price. At the macro level, this implies that domestic producers, when setting their prices, directly take account of foreign prices expressed in domestic currency.

Besides this direct influence, foreign prices also affect domestic prices indirectly through the reaction of nominal wage costs to price changes. Unit wage costs depend on wage cost per employee and on the employment intensity of output. This inverse of apparent labour productivity is itself determined by total factor productivity (the autonomous increase of labour productivity due to technical progress) and by the combination of per capita capital stock and the rate of unemployment, which reflect changes induced by movements in the relative price of factors of production. Wage costs per employee are the result of wage negotiations or of government intervention as a reaction to unfavourable changes in the unemployment rate. In general, wage costs are seen to react to prices, the profit margin of producers, the level of generosity of employment benefits, the tax wedge and the unemployment rate.

Because of the dual relation between price inflation and wage costs, the ultimate causes of price movements are those that directly affect inflation plus all other variables determining wage costs. In this way, foreign inflation, and hence the exchange rate, is a most powerful force that drive domestic prices. This explains why Belgium traditionally chooses exchange rate stability with its most important trading partners as its intermediate target of monetary policy. But even then the inflation differential is not necessarily constant: domestic sources of inflation also remain relevant. More specifically, changes in the unemployment rate in some periods have had a dampening effect on wages and on inflation.

The strong dependence on foreign inflation, often found in a small open economy, implies that the NAIRU, typical of large and relatively closed economies, cannot be derived in the case of Belgium. To obtain insight into the driving forces of the unemployment rate, the demand side also has to be entered into the analysis. The result is that domestic unemployment is determined by both external (the foreign unemployment rate) and internal factors (related to the wage formation process).

Annex 1: Wage-price dynamics in theory

1. Wage bargaining

Wage formation is derived as a simplified version of the model in Layard, Nickell, Jackman (1991),[5] with the union wishing in firm i to maximise utility (income) of potential workers. The objective of the union is thus:

$$u_i = \frac{L_i}{N}\left[\frac{W_i(1-\tau)}{PC}\right] + \left(1 - \frac{L_i}{N}\right)\frac{A_i}{PC} \tag{1}$$

where: u_i = utility;

L = number of employees employed in the firm (this number clearly depends on the outcome of the wage negotiation);

N = potential workers;

W = nominal wage cost;

τ = tax wedge, i.e. the difference between wage cost and net wage;

PC = overall consumer price index;

A = alternative nominal income of an employee outside the firm.

The alternative income consists of two elements:

$$\frac{A_i}{PC} = \left(1 - \varphi_i U_i\right)\frac{\overline{W}_i(1-\tau)}{PC} + \varphi_i U_i \frac{B_i}{PC} \tag{2}$$

where: φU = probability of being unemployed which depends on the unemployment rate scaled with φ;

$\overline{W}(1-\tau)$ = nominal alternative wage that can be earned in other firms;

B = nominal unemployment benefits.

If the wage negotiation is unsuccessful, employees would fall back on utility level u^0, being the alternative income:

$$u_i^0 = \frac{A_i}{PC} \tag{3}$$

Excess income from successful bargaining is, using equations (1) and (3):

$$u_i - u_i^0 = \frac{L_i}{N}\left[\frac{W_i(1-\tau)}{PC} - \frac{A_i}{PC}\right] \tag{4}$$

Firm i, for its part, wishes to maximise profits:

$$\Pi_i = P_i Y_i - W_i L_i - C_i \overline{K}_i \tag{5}$$

[5] Layard, R., S. Nickell and R. Jackman (1991): *Unemployment*, Oxford University Press, Oxford, Chapter 2.

where: Π = profits;

P = output price;

Y = real output;

C = capital cost;

$\overline{K}$ = capital stock, which is a fixed production factor in the short run.

If the wage negotiation fails and workers go on strike, the alternative (negative) profit of the firm is:

$$\Pi_i^0 = -C_i \overline{K_i} \tag{6}$$

The excess profit resulting from a positive bargaining income is derived from equations (5) and (6) as:

$$\Pi_i - \Pi_i^0 = P_i Y_i - W_i L_i \tag{7}$$

The bargaining type which is considered here corresponds to the so called "right to manage" model; union and firm negotiate about the wage level, taking the employment effects into account (meaning that in equations (4) and (7), employment, L_i, should be interpreted as the profit maximisation level of employment). In a second stage, after the wage is fixed, employers decide on the actual level of employment in the firm, which will correspond to its equilibrium value given the negotiated wage level. Using the Nash maximand, the bargaining outcome is the one that maximises:

$$\Omega_i = \left(u_i - u_i^0\right)^{\beta}\left(\Pi_i - \Pi_i^0\right) \tag{8}$$

where: β measures relative union power.

Taking natural logs of (8), differentiating with respect to the gross wage rate, taking into account equations (4) and (7) and noting the envelope theorem $\delta\Pi_i / \delta W_i = -L_i$ (where L should be interpreted as its equilibrium level), the following first-order condition is obtained:

$$\frac{\delta \ln\Omega}{\delta W_i} = \frac{\beta\left[W_i(1-\tau)\frac{\delta L_i}{\delta W_i} + L_i(1-\tau) - A_i\frac{\delta L_i}{\delta W_i}\right]}{L_i\left[W_i(1-\tau) - A_i\right]} - \frac{L_i}{P_i Y_i - W_i L_i} = 0 \tag{9}$$

Working through (9), taking account of equation (2) and aggregating over all firms (supposed to be identical) yields:

$$1 - \frac{B}{W(1-\tau)} = \frac{1}{\varphi U}\left[\frac{-\delta \ln L_i}{\delta \ln W_i} + \frac{1}{\beta\left((PY/WL) - 1\right)}\right]^{-1} \tag{10}$$

Equation (10) determines the aggregate nominal net wage level. It is a positive function of unemployment benefits and a negative function of the unemployment rate, the elasticity of labour with respect to wages (in absolute values) and the labour share. The last two factors may be derived from the producers' optimisation programme.

2. Pricing and labour demand by firms

Production technology is assumed to be Cobb-Douglas:

$$Y_i = TFP_i L_i^{\alpha} K_i^{1-\alpha} \tag{11}$$

where: TFP = index of total factor productivity.

Firms are assumed to operate in markets characterised by monopolistic competition. The demand for their products is drawn from an "Almost Ideal Demand System" (AIDS), such that the market share of an individual firm is given by:

$$S_i = \frac{P_i Y_i}{R} = \omega_i + \sum_i \gamma_{ij} \ln P_j + \chi_i \ln \frac{R}{IP} \tag{12}$$

where: R = total nominal demand;

IP = general aggregate index of prices;

and with:

$$\ln IP = \omega_0 + \sum_k \omega_k \ln P_k + ½ \sum_k \sum_l \gamma_{kl} \ln P_k \ln P_l$$

and:

$$\gamma_{kl} = \gamma_{lk}, \qquad \sum_k \omega_k = 1, \qquad \sum_k \gamma_{kl} = 0, \qquad \sum_k \chi_k = 0.$$

The firm's optimum is obtained by maximising:

$$P_i Y_i - W_i L_i - C_i \overline{K_i}$$

subject to equations (11) and (12).

After substitution, the objective function can be rewritten as:

$$\max.\psi = \left[\omega_i + \sum_j \gamma_{ij} \ln P_j + \chi_i \ln \frac{R}{IP}\right] R - W_i L_i - C_i \overline{K_i} - \lambda \left[\left(\omega_i + \sum_j \gamma_{ij} \ln P_j + \chi_i \ln \frac{R}{IP}\right) \frac{R}{P_i} - TFP_i L_i^{\alpha} \overline{K_i^{1-\alpha}}\right]$$

Maximisation with respect to λ, P_i and L_i yields the following first-order conditions (under the assumption that IP, R and competitors' prices are given for each individual firm):

$$P_i = \frac{S_i R}{TFP_i L_i^{\alpha} K_i^{1-\alpha}} \tag{13}$$

$$\lambda = \left(\frac{\gamma_{ii}}{\gamma_{ii} - S_i}\right) P_i \tag{14}$$

$$W_i = \lambda TFP_i \alpha L_i^{\alpha-1} \overline{K_i^{1-\alpha}} \tag{15}$$

where, in terms of the absolute price elasticity of demand (η),

$$\left(\frac{\gamma_{ii}}{\gamma_{ii}-S_i}\right)=1-\frac{1}{\eta_{ii}} \qquad \text{with } \gamma_{ii} \text{ supposed to be} < 0 \text{ such that } \eta_{ii} > 1.$$

The absolute price elasticity is a negative function of the market share.

Converting equations (12) - (15) into growth rates gives the following system of equations:

$$\overset{\circ}{P_i} = \overset{\circ}{S_i} + \overset{\circ}{R} - \overset{\circ}{TFP_i} - \alpha \overset{\circ}{L_i} \tag{16}$$

$$\overset{\circ}{\lambda_i} = \overset{\circ}{P_i} - \frac{S_i}{S_i - \gamma_{ii}} \overset{\circ}{S_i} \tag{17}$$

$$\overset{\circ}{L_i} = -\frac{1}{1-\alpha}\left[\overset{\circ}{W_i}\right] + \frac{1}{1-\alpha}\overset{\circ}{\lambda} + \frac{1}{1-\alpha}\overset{\circ}{TFP_i} \tag{18}$$

$$\overset{\circ}{S_i} = \sum_j \frac{\gamma_{ij}}{S_i}\overset{\circ}{P_{ij}} + \frac{\chi_i}{S_i}\overset{\circ}{R} - \frac{\chi_i}{S_i}\overset{\circ}{IP} \tag{19}$$

From the reduced-form solution, it can be shown that:

$$\left.\frac{\delta \ln L_i}{\delta \ln W_i}\right|_{IP,R,P_j} = \frac{-(S_i - \gamma_{ii})^2}{(S_i - \gamma_{ii})^2 - \alpha\gamma_{ii}^2} < 0 \tag{20}$$

from which it follows that the absolute wage cost elasticity of labour demand is decreasing in S_i. When aggregating over n identical domestic firms, the demand system boils down to the allocation of total consumption between domestic (i) and foreign goods (j):

$$S_i = \frac{PY}{nR} = \omega_i + n^* \gamma_{ij} \ln\frac{P^*}{P} + \chi_i \ln\frac{R}{IP} \tag{21}$$

where: P^* = foreign price level;

n^* = number of foreign firms j;

$\gamma_{ij} > 0$; that is, we expect the aggregate market share to be a positive function of the foreign price deflated by the price of domestic firms (the effective exchange rate).

From equations (14) and (15) follows:

$$\frac{W_i L_i}{P_i Y_i} = \frac{WL}{PY} = \frac{-\alpha\gamma_{ii}}{S_i - \gamma_{ii}} \tag{22}$$

which can also be written as:

$$P = \frac{(\gamma_{ii} - S_i)}{\gamma_{ii}} \frac{WL}{\alpha Y} \tag{23}$$

It follows that although we assumed a Cobb-Douglas production technology, the labour share and price mark-ups are non-constant. Mark-ups tend to increase and the labour share tends to decline when the market share improves because firms' price elasticity of demand is inversely related to their market share.

Substituting equations (20) and (22) in (10) yields the following wage equation:

$$1 - \frac{B}{W(1-\tau)} = \frac{1}{\varphi U}\left[\frac{(S_i - \gamma_{ii})^2}{(S_i - \gamma_{ii})^2 - \alpha\gamma_{ii}^2} + \frac{\alpha\gamma_{ii}}{\beta[(S_i - \gamma_{ii}) + \alpha\gamma_{ii}]}\right]^{-1} \tag{24}$$

with: $\frac{\delta W}{\delta S_i} > 0.$

Both the labour share and the wage elasticity of labour depress the wage level and both of them are decreasing in the firm's market share. Hence, W is positively influenced by the domestic market share or the relative price of competitors. This result is independent of any myopia on the part of wage earners. Foreign prices raise the domestic wage because the wage outcome is positively affected by the lower wage elasticity of labour and the lower labour share that a rising market share implies.

The unemployment rate U can be defined as:

$$U = 1 - \frac{L}{L^S} \tag{25}$$

where: L^S = labour supply.

The complete model now consists of equations (11), (21), (23), (24) and (25). Linearising these equations, making some substitutions and assuming χ_i = 0, the following price and wage equations can be obtained:

$$\ln P = \pi_0 + \pi_1 S_i + \frac{1-\alpha}{\alpha}\left(\ln\frac{L^S}{K} - U\right) - \ln TFP + \ln W \tag{26}$$

or alternatively:

$$\ln P = \pi_0 + \pi_1 S_i + \frac{1-\alpha}{\alpha}\ln\frac{Y}{K} - \frac{1}{\alpha}\ln TFP + \ln W \tag{26'}$$

$$\ln W = \sigma_0 + \ln\frac{B}{(1-\tau)} - \sigma_1 U + \sigma_2 S_i \tag{27}$$

with:

$$S_i = \rho_0 + \rho_1 \ln\frac{P^*}{P}. \tag{28}$$

In Layard-Nickell type models, given the real unemployment benefits, the equivalents of (26) and (27) are enough to derive an equilibrium unemployment rate. Here, the effective exchange rate is still present and the demand side of the model is necessary to derive the equilibrium.

Annex 2: Estimations

The wage-price model from Appendix 1 was estimated on the period 1963-94. Equations (26') and (27) are used with S_i replaced according to (28), giving the following long-run equations:

$$\ln P = a_0 + a_1 \ln P^* + (1-a_1)\ln W + a_2 \ln\frac{Y}{KC} + a_3 \ln\frac{OIL}{P}$$

$$\ln\frac{W}{P} = b_0 + b_1 \ln\frac{B}{P} + b_2 \ln(1-\tau) + b_3 U + b_4 \ln\frac{P^*}{P} + b_5 \ln\frac{Y}{KC} + b_6 \ln\frac{OIL}{P}$$

The specifications are more general than the theoretical model:

- on the basis of preliminary results, a relative oil price seems necessary to explain the Belgian mark-up. Since the wage bargaining anticipates the price formation process, the oil price is also introduced in the wage equation.
- the parameter on productivity in the price equation is independent of the wage coefficient (productivity plays through *Y/KC*, a combination of *Y/K* and *TFP* which, in logarithm, is a multiple of labour productivity.
- the impact of unemployment benefits on wages is not constrained to unity and productivity was introduced to conform to traditional specifications of wage formation.[6]
- the tax wedge impact is not constrained to -1 which would imply that all increases in direct taxes or social security contributions would rest on companies.

Long-run equations are introduced in an ECM-type specification to take dynamic adjustment into account:

$$\begin{aligned}\Delta\ln P = {} & c_0 + c_1\Delta\ln P(-1) + c_2\Delta\ln P^* + c_3\Delta\ln P^*(-1) + c_4\Delta\ln W + c_5\Delta\ln W(-1) \\ & + c_6\Delta\ln OIL + c_7\Delta\ln OIL(-1) + c_8\Delta\ln\left(\frac{Y}{KC}\right) + c_9\Delta\ln\left(\frac{Y}{KC}\right)(-1) \\ & + c_{10}\Delta\ln e + c_{11}\Delta\ln e(-1) + c_{12}DUC + c_{13}DUC(-1)\end{aligned}$$

6 Extensions to the theoretical model would also justify this addition.

$$-\mu\left(\ln P - a_1 \ln P^* - (1-a_1)\ln W - a_2 \ln\frac{Y}{KC} - a_3 \ln\frac{OIL}{P}\right)(-2)$$

Two more variables are introduced in the dynamic part, exchange rate differences to allow for incomplete pass through in the price formation and a degree of capacity utilisation (*DUC*) which is the divergence between the potential output level attainable on the basis of currently *hired* factors and realised output. Such a GDP gap is common in NAIRU-type approaches.

$$\begin{aligned}\Delta\ln W = {} & d_0 + d_1\Delta\ln W(-1) + d_2\Delta\ln P + d_3\Delta\ln P(-1) + d_4\Delta\ln P^* + d_5\Delta\ln P^*(-1)\\ & + d_6\Delta\ln OIL + d_7\Delta\ln OIL(-1) + d_8\Delta\ln B + d_9\Delta\ln B(-1)\\ & + d_{10}\Delta\ln(1-\tau) + d_{11}\Delta\ln(1-\tau(-1)) + d_{12}\Delta U + d_{13}\Delta U(-1)\\ & + d_{14}\Delta\ln\left(\frac{Y}{KC}\right) + d_{15}\Delta\ln\left(\frac{Y}{KC}\right)(-1)\\ & -\upsilon\left(\ln\frac{W}{P} - b_1\ln\frac{B}{P} - b_2\ln(1-\tau) - b_3U - b_4\ln\frac{P^*}{P} - b_5\ln\frac{Y}{KC} - b_6\ln\frac{OIL}{P}\right)(-2)\end{aligned}$$

The 3SLS results are:[7]

$$\begin{aligned}\Delta\ln P = {} & 0.65 + 0.56\Delta\ln P^* + 0.39\Delta\ln W(-1) + 0.01\Delta\ln OIL - 0.01\Delta\ln OIL(-1)\\ & + 0.56\Delta\ln e - 0.16\Delta\ln e(-1) + \left(0.18\ln\frac{P^*}{P} + 0.25\ln\frac{W}{P} + 0.11\ln\frac{Y}{KC} - 0.02\ln\frac{OIL}{P}\right)(-2)\end{aligned}$$

DW = 2.21 s.e. = 0.007.

$$\begin{aligned}\Delta\ln W = {} & -0.45 - 0.99\Delta\ln W(-1) + 0.33\Delta\ln P(-1) + 0.23\Delta\ln P^* + 0.32\Delta\ln P^*(-1)\\ & + 0.01\Delta\ln OIL - 0.01\Delta\ln OIL(-1) + 0.26\Delta\ln B + 0.28\Delta\ln B(-1)\\ & - 0.78\Delta\ln(1-\tau) - 0.89\Delta\ln(1-\tau(-1)) - 0.01\Delta U(-1)\\ & -\left(0.73\ln\frac{W}{P} - 0.46\ln\frac{B}{P} + 0.89\ln(1-\tau) + 0.019U - 0.23\ln\frac{P^*}{P}\right)(-2)\end{aligned}$$

DW = 2.00 s.e. = 0.01.

Nearly all the coefficients are significant at the 5% level or less; only the coefficient on Δ ln*P* in the wage equation is significant at the 10% level. We can make a few observations:

- our wage model looks quite plausible. All suggested variables are present and productivity was not significant. Thus, there is a relationship between the wage level and the unemployment rate. NAWRU-type approaches cannot be justified for Belgium since they ignore the levels in both wages and prices;
- oil prices have an impact on the long-run mark-up and on the dynamics of both equations but not on long-run wage formation;

---

[7] *B*, *U* and *Y* are taken as (potentially) endogenous, with past values used as instruments.

- in the long run, the coefficient on *Y/KC* in the price equation is such that the coefficient on labour productivity would equal -0.27, close in absolute value to the wage coefficient. Hence, it is the unit labour cost that is present in the price formation;
- adjustments to exchange rate changes are sluggish. The exchange rate and the foreign price share the same coefficient in the current period. It is thus the foreign price in foreign currency that is relevant in the short run;
- the relative influences in the price formation change with time: from foreign prices (competition) in the short run to domestics costs, that is wages, in the long run;
- our GDP gap measure turned to be nonsignificant, which puts into question NAIRU-type approaches in addition to the fact that we have an explanation relating the price levels to the unemployment rate in the reduced form.

To recover the steady-state equations for price and wage inflation used in Sections 1 and 2, one simply takes the difference of the long-run part of the dynamic equations (the following identity derived from the production function is also used: $\Delta\ln(Y/KC) = -.7/.3\ \Delta\ln\ TFP + .7\Delta\ln(L^S/K) - .7\Delta(U/100)$, where .7 is the labour elasticity of output and U is in percent).

Annex 3: General equilibrium

The reduced-form wage-price model gives a link between Belgian inflation in the long run and unemployment rate variations as well as other factors. It was introduced in Section 3 and can be rewritten:

$$\left(\frac{\overset{\circ}{P}}{P*}\right) = 3.12\left(\frac{\Delta UD - \Delta U}{100}\right)$$

where:

$$\frac{\Delta UD}{100} = \frac{-0.10\left(\frac{\overset{\circ}{OIL}}{P*}\right) + 0.67\overset{\circ}{b} - 1.32\left(1-\overset{\circ}{\tau}\right) - 1.13\overset{\circ}{TFP} - 0.34\left(\frac{\overset{\circ}{K}}{L^S}\right)}{3.12}$$

$$= -0.03\left(\frac{\overset{\circ}{OIL}}{P*}\right) + 0.22b - 0.42\left(1-\overset{\circ}{\tau}\right) - 0.36\overset{\circ}{TFP} - 0.11\left(\frac{\overset{\circ}{K}}{L^S}\right)$$

ΔUD accounts for the factors other than unemployment affecting the inflation differential and the relation can be seen as a supply curve: a higher activity level (lower unemployment) requires higher prices. It is important to recognise that an opposite relationship should exist, a "demand" curve, with a negative link between prices (low competitiveness) and demand, that is a positive link between prices and unemployment.

We don't pretend to model, in this paper, the complicated relationship between competitiveness and unemployment. But to shed some light on the impact of this second relation on the link between inflation and unemployment, we make a crude approximation. An equation relating the variations in the unemployment rate differential (Belgium with respect to the EU) to our inflation differential was estimated and gave the following long-run demand equation (in variations):

$$\left(\frac{\Delta U - \Delta U^*}{100}\right) = 0.061\left(\frac{\overset{\circ}{P}}{P^*}\right) = \frac{1}{16.5}\left(\frac{\overset{\circ}{P}}{P^*}\right)$$

No significant constant was found implying a continuous divergence of Belgian unemployment from its EU counterpart. Combining supply and demand, we get

$$(\Delta U - \Delta U^*) = \frac{3.1}{3.1+16.5}(\Delta UD - \Delta U^*)$$

and

$$\left(\frac{\overset{\circ}{P}}{P^*}\right) = \frac{16.5*3.1}{3.1+16.5}\left(\frac{\Delta UD - \Delta U^*}{100}\right)$$

The same factor is present in the unemployment and inflation differentials, the difference between exogenous inflationary factors and unemployment evolution abroad. One could speak of a domestic component (*OIL/P** excepted) and a foreign component. This is even more apparent if we rewrite ΔU as

$$\Delta U = \frac{3.1}{3.1+16.5}\Delta UD + \frac{16.5}{3.1+16.5}\Delta U^*$$

where, relying on relative weights, the foreign component seems the most important. The 3.1 coefficient is a multiple (via the wage-price spiral) of the sensibility coefficient of wages to the unemployment rate; if this reaction increases, thanks to the denominator, the domestic as well as the foreign factors will loose some impact and the unemployment rate changes will be lower. What are the implications of the demand curve for our previous analysis of inflation? The final price equation stays close to the previous one since the 16.5/(3.1+16.5) factor is close to unity. But it is the foreign unemployment rate that now appears in the equation, not the domestic one. Their evolutions are, however, similar. Finally, the presence of the same factor in the unemployment and inflation differential suggests that the same change in domestic factors could lower unemployment and inflation at the same time. The fights against inflation and unemployment would be compatible in the long run. However, this conclusion relies upon our demand approximation and upon the fact that this change would not immediately affect domestic activity.

Comments on: "Inflation and unemployment in Belgium" by Michel Dombrecht and Philippe Moës

by Gregory Sutton

This paper represents a significant theoretical advance in the area of inflation determination in small open economies. From their model, we learn that the NAIRU concept may not be a particularly useful one for understanding the behaviour of inflation in small open economies. For these economies, foreign inflation may be a more important determinant of domestic inflation than labour market conditions. The authors use their model to interpret the Belgian inflation history and conclude that foreign inflation was the principle cause of domestic inflation over the 1962-94 period.

From a theoretical perspective, an important contribution of the paper is the incorporation of imperfect competition into an equilibrium model of price and wage determination. By relaxing the assumption of perfect competition in product markets, the authors are able to analyse the impact of foreign inflation on domestic inflation that arises from the equilibrium response of profit margins to shifts in consumer expenditure between domestic and foreign goods. A decline in foreign prices, holding domestic prices and the exchange rate constant, induces a shift of consumption expenditure from domestically produced goods to foreign goods. This leads to a decline in the market shares of domestic firms, inducing them to lower profit margins and prices. Likewise, an increase in foreign prices works to raise domestic prices. In this way, foreign inflation has a direct impact on domestic inflation via the price-setting behaviour of domestic firms.

As noted by the authors, there are also indirect influences of foreign inflation on domestic inflation. These include the impacts of import prices on the outcome of wage negotiations and on production costs.

Chart 7 of the paper plots both the profit margin and the relative GDP deflator in Belgium. There is clearly a positive relation between changes in these series, as predicted by the model. The results of formally estimating the model are also encouraging. The evolution of the rate of inflation predicted by the model closely follows actual Belgian inflation over the period studied. The conclusion reached from the empirical exercise is that foreign inflation was the principle cause of Belgian inflation over the sample period, with the growth rate of per capita capital stock and the slowdown in productivity also contributing significantly.

Therefore, I believe that the authors have made an important theoretical contribution to our understanding of the causes of inflation in small open economies. The concept of equilibrium unemployment employed in the paper, a level of unemployment that equates foreign and domestic inflation rates, is certainly a more reasonable measure than the NAIRU for the case of a small open economy.

NAIRU, incomes policy and inflation

Silvia Fabiani, Alberto Locarno, Gian Paolo Oneto and Paolo Sestito[1]

Introduction

Italy has been one of the few industrialised countries resorting to incomes policies in the current decade. Starting from 1992, the previous wage indexation mechanism (the *scala mobile*) has been dismantled and the institutional framework of wage negotiation has been reshaped with the aim of strengthening the link of nominal wages to target inflation. Furthermore, the role of the incomes policy agreements has been pervasive as they have implied a continuous involvement of social partners in the implementation of economic policy (the so-called *concertazione*), particularly in the definition of the fiscal and social policy design.

Many observers have claimed that the typical target of traditional incomes policies has been attained, for a remarkable slowdown in wage and price inflation has been taking place since 1992 despite the inflationary impulses due to two episodes of sharp depreciation of the exchange rate (Figure 1).

However, no apparent progress has been made concerning labour market imbalances: in 1996 the unemployment rate exceeded by 4 percentage points the already high level – largely reflecting structural factors – inherited from the 1980s (Figure 2).

The rise in unemployment and the slowdown in GDP growth experienced since 1992 may have been more important than the incomes policy episode in moderating wage growth and inflation; more generally, economic policy and, in particular monetary policy, has become much more committed to price stability.

The various analytical issues raised by this experience may be summarised as follows:

i) How important was the wage moderation engineered by the incomes policy framework in checking inflationary pressures?

ii) Is there any evidence of a long-run change in the bargaining structure and the wage formation mechanism? In particular, are there signals that the new bargaining system may lead to a lower NAIRU? Are there implications for economic policy stemming from the changes in the dynamic adjustment of nominal wages to prices?

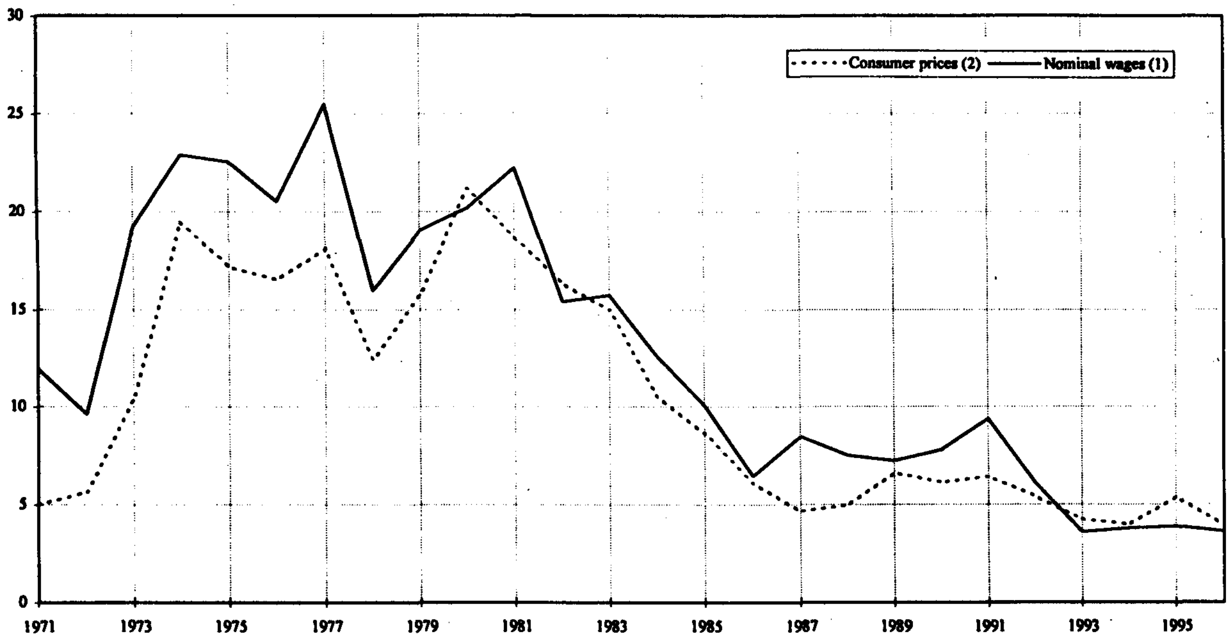
iii) Is there any relationship between the incomes policy framework and other facets of the design of economic policy, namely the fiscal stance and the reshaping of labour market regulations and social policy institutions? In particular, has unions' involvement (the *concertazione*) favoured fiscal adjustment, providing the necessary consensus in a period of extreme political turmoil, or has it hindered the process of structural reforms?

Our answers to these sets of questions are preliminary and partial. First, we do not deal with the third, most difficult, set of questions. As far as the second question is concerned we provide only a partial answer. We analyse the possibility of structural breaks in the wage formation process using a standard Phillips-curve specification embodied in a large macroeconometric model (the Bank

1 S. Fabiani, A. Locarno and P. Sestito are members of the Research Department, Bank of Italy, and G. P. Oneto of ISCO, the National Institute for the Study of the Economics Situation. This paper has benefited from comments received during presentations in Rome (Bank of Italy), Basel (BIS) and the University of Milan; however, the authors are the only responsible for the opinions expressed, which do not necessarily involve the Institutions to which they belong.

Figure 1

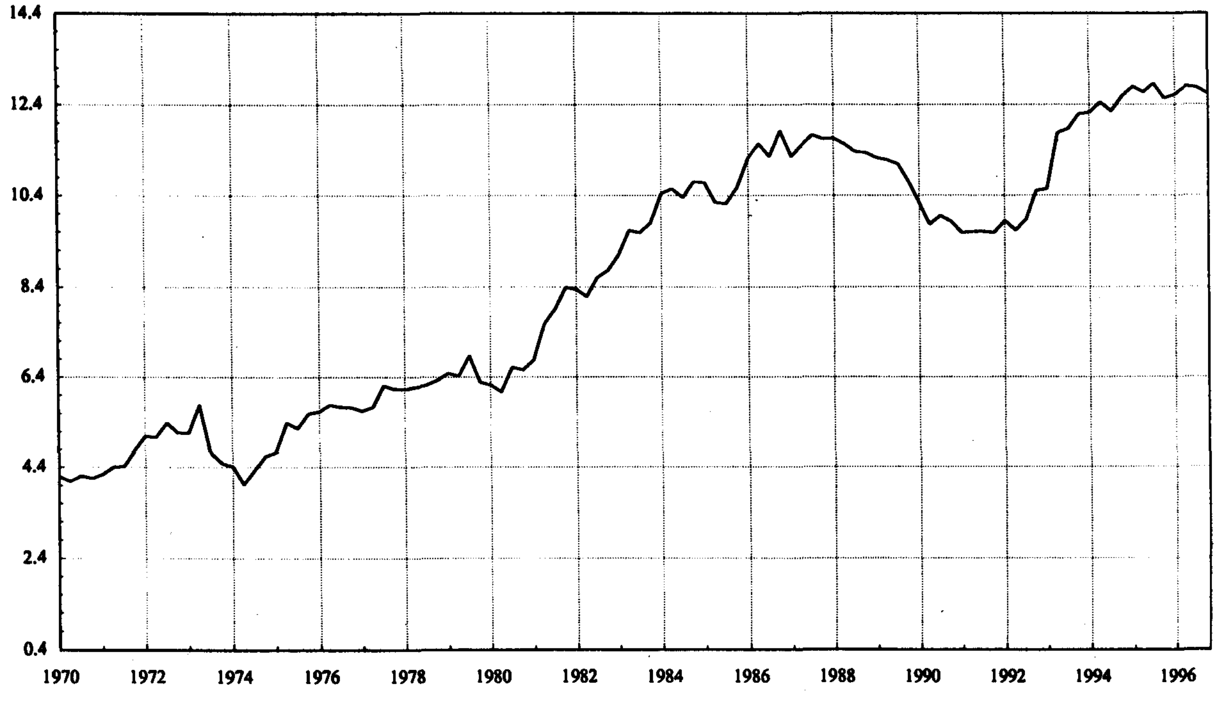
Price and wage inflation
Yearly percentage changes



[1] Per capita earnings in the non-farm non-energy private sector. [2] Cost-of-living index.

Figure 2

Unemployment rate*
Yearly percentage changes



* Including workers paid by the Wage Supplementation Fund. Data before 1992 were made compatible with the current definition (see Casavola (1994)).

of Italy quarterly model, henceforth the BIQM). We consider both the hints of shifts in the NAIRU implicitly included in the Phillips curve and the changes in the dynamic adjustment of wages to prices; moreover, embedding this equation into a large model we try to identify the overall[2] effects of the wage moderation episode. However, we believe that the specification of the labour market adopted in the BIQM is not the most suitable one for identifying the role of the structural factors highlighted in the literature, both for Italy and other industrialised countries, as determinants of the NAIRU (and of the sluggishness of the unemployment rate in returning to the NAIRU). We are currently working along this specific route, trying to identify a variable measure of the NAIRU.[3]

Two additional *caveats* must be stressed at the outset. First, our analysis of the structural break deriving from the reshaping of the bargaining system is an early evaluation, given that only few years have elapsed since the reform was agreed upon and many aspects of the new regime, if anything really new is emerging, have still to be defined. Moreover, our analysis is based on an aggregate wage equation which may not be the most suitable analytical tool for identifying the main features of the dynamic process of wage determination (which includes a wage-to-wage process on top of the wage-to-price adjustment).

The paper is organised along the following lines. In the next section we describe the Italian bargaining system. In Section 2 we focus on the recent incomes policy experience and the novelties it introduced into that system. We present in Section 3 a comprehensive econometric analysis of the stability of the wage equation included in the BIQM. Having identified some evidence of breaks, we try to characterise the new wage determination mechanism (Section 4), and to evaluate the overall macro effects of these changes resorting to simulations of the whole model (Section 5). Finally, we present some overall conclusions, focusing on the features of the new bargaining system that remain unsettled.

1. The Italian wage bargaining system

Two main issues have been emphasised in the Italian debate on wage setting: the effects on inflation persistence stemming from the automatic wage indexation mechanism (the *scala mobile*); the rigidity wages (both the average real wage and wage differentials) with respect to labour market unbalances.[4]

In the bargaining framework fully effective up to the end of 1991, wages resulted from the adding up of pay increases determined by the automatic indexation mechanism and the effects of

---

2 As said, excluding the possible interactions between incomes policy and the fiscal stance.

3 In a related research project we are trying to identify and measure the long-run structural unemployment rate. On practical grounds we are implementing both simple S-VAR systems à la Blanchard-Quah, identifying the shocks to the NAIRU separately from demand and supply shocks which should have only temporary effects on unemployment, and a reduced-form specification for the unemployment rate including both cyclical and structural determinants of unemployment and, in particular, those related to the wage bargaining process. The analysis is carried out for a longer period, including both the fall of unemployment in the 1950s and its subsequent rise from the beginning of the 1970s. In the forthcoming paper we also plan to compare our procedure with the standard practice of looking directly at the unemployment-(accelerating) inflation trade-off popularised by Elmeskov (1993) and refined, among others, by Staiger, Stock and Watson (1996).

4 For a general overview of the structural aspects of the inflationary process in Italy, including those related to the wage bargaining system, see Visco (1994). For a previous study using (an earlier vintage of) the BIQM for characterising the NAIRU see Bodo and Visco (1987); Gavosto et al. (1995) use the BIQM for examining the stability of the NAIRU. A survey of the last two decades of empirical specifications of the wage curve in Italy is in Sestito (1994). The aspect of the rigidity in the wage structure mostly emphasised is the one concerning the regional differentials (see Bodo and Sestito (1991), Faini (1995) and Casavola et al. (1995)).

two separate levels of bargaining. These were the national industry-level contracts lasting three years and bearing on almost every aspect of the wage structure and work organisation, and an additional company-level bargaining taking place mainly in large firms.

The indexation provision established an automatic link between the evolution of wages and past price changes, covering all employees. Such a mechanism strengthened workers' bargaining power, reinforcing real wage resistance to external shocks (e.g. the two oil shocks in 1973-74 and 1979) and changes in indirect taxation; moreover, it strengthened nominal inertia in wage growth, causing above all a strong persistence of the effects of exogenous inflationary shocks. During the 1980s, the functioning of the mechanism underwent a number of changes that lowered, over time, the degree of indexation and its compressing effects on wage differentials, and which were negotiated at the central level and were favoured by a more cooperative stance in labour relations. In the early 1980s, the actual elasticity to CPI guaranteed by the scheme was on average close to 0.8; adjustments were taking place quarterly and the mechanism produced a significant compression of differentials (the same absolute increase was granted to everybody).[5] The degree of indexation was reduced between 1982 and 1984 through ad hoc modifications agreed to in centralised negotiations that involved the social partners and the Government (the so-called Tripartite Agreements). A new indexation clause was introduced in 1986: the degree of coverage with respect to CPI inflation was stabilised around 0.5-0.6, with a half-year adjustment lag and a milder compression of wage differentials.

The overwhelming role played by centralised elements (the indexation component and the industry-level settlements) is thought to have produced a wage structure characterised by a low degree of differentiation and flexibility.[6] Wages were not allowed to adjust to movements in productivity or in relative labour demand occurring in specific segments of the labour market (in particular at the regional level). The absence of coordination between the different bargaining levels is deemed to have resulted in an inflationary bias, strengthened by the backward-looking indexation mechanism. National contracts were rather independent across industries, whereas a clear pattern of bargaining did not emerge and firm-level agreements established pay increases resulting in widespread wage drifts. Finally, the backward looking indexation (effective only vis-à-vis price rises) prolonged over time the effects of inflationary shocks. All in all, many observers have considered Italy as characterised, in the Calmfors-Driffil (1988) framework, by an intermediate degree of centralisation, with negative externalities arising from uncoordinated wage pressures and rigidities stemming from the lack of responsiveness to "local" market conditions.[7]

However, ideological and politically based motivations influencing the behaviour of the trade unions have acted as a centralising device. The political leadership of unions is thought to have sometimes enhanced policies of coordination and explicit wage moderation. By and large, these phases stemmed from situations characterised by economic turmoil calling for "emergency" interventions. A first episode was at the end of the 1970s, when a mild form of moderation in wage

---

5 Until the mid 1970s the absolute rises granted by the *scala mobile* were partly differentiated across sectors and pay levels; an agreement set in 1975 after bitter conflicts equalised (at the highest amount) the absolute rises. As a result, the actual degree of indexation for an average worker (in the manufacturing sector) rose from about 0.60 to about 0.90 in 1977; a small decline (to 0.80) took place in the following years as the mechanism did not take into account the drift in real wages. For details on the *scala mobile* mechanism see Banca d'Italia (1986). A reconstruction of the degree of indexation during the 1950s and 1960s is in Robotti (1973). An evaluation of the "optimality", in the classical sense of Gray (1976), is in De Stefanis (1995).

6 Formally, only firms belonging to the employers associations which signed the contracts are obliged to apply them to all their employees. In practice these are applied in the whole "regular" sector, both because the courts tend to refer to the national contracts in case of labour disputes and because several pieces of regulations prevent the eligibility for some incentives for firms that do not apply national-contract provisions. A proxy measure of coverage may, therefore, be obtained by considering the share of workers employed in the "regular" sector as defined in the national accounts.

7 See, for instance, CEC (1993) and Demekas (1995).

claims was introduced through an agreement involving the major unions.[8] More troubled was the attempt made during 1982-84, as the biggest union finally withdrew from the incomes policy framework. In general, a more cooperative stance of unions is deemed to have emerged during the 1980s, as the phase of social conflicts that started in 1968-69 came to an end.

Probably the most important experience of coordination is the one that took place in 1992 and 1993 when wage setting was defined centrally by a sequence of agreements between the social partners, that ended with the introduction in July 1993 of a new mechanism of wage negotiation embodying incomes policy guidelines. The agreements were shaped by the growing concern about the deterioration of the economic situation. At the beginning of the 1990s, the effects of maintaining the exchange rate parities in the ERM, together with a growth in nominal wages higher than in the other European countries, were undermining the competitiveness of manufacturing firms. Consumer price inflation was fuelled by a still buoyant economic activity in the service sector and the other sheltered industries, favoured by the demand effects of a rising public debt. The perception of a risk of fiscal unsustainability was gradually emerging; in 1992 the lira was one of the first currencies being put under strong pressure when the ERM crisis started.

At the end of 1991, when employment was beginning to decline, a first agreement established a temporary freeze of the wage indexation mechanism. Then, in July 1992, the social partners agreed upon the complete dismantling of the *scala mobile* and suspended firm-level bargaining up to the end of 1993, in exchange for an across-the-board rise (a flat-rate bonus worth about 0.7% of the average monthly compensation) to be paid in January 1993. Notwithstanding the agreement, in the autumn of 1992 unprecedented economic and financial turmoil hit Italy. In September the lira was forced to leave the ERM; in December the devaluation reached 15%. The Government announced a very large fiscal package, including a reform of the pension system and a freeze of public sector wages; the budget package was approved in December. By that time the economy had entered a severe recession which aggravated employment losses.

The final agreement reached in July 1993 established the institutional features of a new bargaining framework aimed at introducing two main elements: a better coordination of the different components contributing to wage evolution and a pivotal role of future (target) inflation in shaping the development of nominal wages.

2. The new bargaining framework

The system established in July 1993 has maintained a two-tier configuration. The national industry-level contracts set the structure and evolution of wages (over a two-year horizon) as well as working conditions (with four-year validity).[9] According to the general guidelines spelled out in the agreement, pay rises set by industry-level contracts should be "consistent" with the official inflation targets. Unexpected inflation does not result in automatic wage rises,[10] but the differential between actual and target inflation in the past two years is one of the aspects to be considered at the renewal of sectoral wage settlements (taking into account terms-of-trade movements and the evolution of actual wages). In turn, the decentralised bargaining, which takes place mostly in large firms, should

8 For an analysis of the relationships between industrial reorganisation and changes in industrial relations in Italy since the 1970s, see Barca and Magnani (1989). Giavazzi and Spaventa (1989) compare the Italian and the UK experiences.

9 They also establish general guidelines by sector for firm-level negotiations; whether a specific firm is to be covered by decentralised bargaining still depends on local conditions, in particular unions' strength.

10 An indexation mechanism is maintained only in the case of a significant delay in the renewal of national contracts: three (six) months after expiration, the wage levels set by the old contract are increased by 30% (50%) of the target inflation rate.

aim at strengthening the flexibility of wages with respect to firms' performance, linking pay bonuses to specific productivity or profitability indicators.

Starting from the debate on corporatism, a large amount of literature has analysed the relationship between the bargaining structure and macroeconomic performance. Several aspects of the bargaining structure have been considered but the one most emphasised has been the degree of centralisation. Calmfors and Driffill (1988) have claimed that there is an hump-shaped relationship between the degree of centralisation and equilibrium unemployment. However, Calmfors (1993) has shown that several facets of a bargaining system may not be easily synthesised by a single variable meant to measure the degree of centralisation. For instance, decentralised bargaining units may be strongly coordinated through guidelines set at a central level, while formally centralised systems may have mixed features deriving from the effects of widespread wage drift and plant level bargaining. As argued by other authors (see Layard et al. (1991)), the Calmfors and Driffill hump-shaped relationship may reflect two separate factors: "coordination" and "unions' strength", respectively with negative and positive effect on the NAIRU and usually both positively correlated with the degree of centralisation. Even more important is the fact that the shape of the relationship between unemployment (and other macroeconomic performance indicators) and the degree of centralisation has not always been confirmed in the empirical studies which have followed.[11] Actually, this instability of the relationship may derive from the interaction between the bargaining structure and the nature of the shocks experienced by an economy: a centralised structure may be better suited to offset aggregate and undifferentiated shocks, while a more decentralised one may more promptly respond to "structural" and micro-based shocks. The empirical estimates may lead to robust results only if appropriate variables controlling for the nature of the shocks are introduced.

It is therefore not correct to analyse the changes determined by the new system by focusing uniquely on the degree of centralisation. A more complete analysis of its features is needed for understanding the possible effects of changes in the wage-setting procedures.[12]

A first element is related to the increase in coordination possibly arising from the link with target inflation, which should act as a common guideline for all sectors. Moreover, coordination might be strengthened if a clearer distinction emerges between the domains of industry and firm-level bargaining. What is not clear is whether the incomes policy agreements have led to relevant changes in the actual degree of coordination. Both in 1993-94 and in 1995-96, all national industry-level contracts have been settled with the inflation target set by the Government (see Irs (1997)) maintained as a reference. However, no trend towards a more synchronised structure has emerged. Moreover, the apparently strong coordination across bargaining units obtained so far might have been a temporary feature determined by the cooperative behaviour adopted by unions to deal with the "emergency" in the economic and political situation. Furthermore, the distinction between the domains of firm and industry-level bargaining is still an open issue.

As for other aspects of the sensitivity of real wages to labour market disequilibria, the overall effect of the changes brought about by the reform is even more difficult to evaluate. On the one hand, the framework established in the incomes policy agreements has been interpreted by some observers as a further step towards a more cooperative stance in industrial relations,[13] possibly strengthening the effect of unemployment in moderating real wages.[14] On the other hand, the

---

[11] See OECD (1997).

[12] An initial discussion of the possible novelties implied by the new bargaining framework is in Dell'Aringa (1995).

[13] See, among others, Regini and Regalia (1996), in which the move towards a co-operative stance at the national level is viewed as only partly due to the emergency phase (the need for fiscal adjustment and the exchange rate crisis, on the one hand, and the crisis of the traditional political parties, on the other). The move has also been related to the co-operative experiences developed at the firm level during the 1980s.

[14] As usual it is quite difficult to distinguish between changes in unions' preferences and variations in unions' power; an empirical evaluation for Italy has been presented in Checchi (1995).

centralisation of bargaining so far implicit in the implementation of the incomes policy agreements may impair the ability to produce the differentiation in the wage structure needed to counteract the labour market segmentation plaguing Italy.

An additional aspect to be considered is the abolition of the old indexation mechanism with its distortive effects (on wage differentials and on real wage rigidity with respect to changes in the terms of trade and indirect taxes).

All these changes may have an impact on the determination of labour market equilibrium, possibly resulting in a shift of the unemployment rate consistent with stable wage inflation (i.e. the NAIRU). The new system might also lead to a change in the dynamic pattern of the adjustment of wages to prices and, therefore, to a change in the inflationary process. The absence of indexation and the two-year horizon of the national contracts are likely to increase the inertia of nominal wages (in levels), slowing down the adjustment and increasing the weight of expected inflation in the determination of wages. For a given level of expected inflation, the short-run impact of inflationary shocks should be reduced. These shocks will somehow feed back later on, when re-negotiation occurs; in this sense the absence of indexation, coupled with a quite long duration of contracts (two years), may produce fluctuations in real wages unrelated to the fundamentals of the economy and a longer memory in the inflation process.

Assuming as a reference model the standard long-run vertical Phillips curve, the changes in the pattern of nominal adjustment should have no implications for the NAIRU. However, in the short run there might be "real" effects. In the theoretical literature, wage indexation is beneficial for disinflation because it implies that wage setters do not need to discount the risk of monetary shocks. The cost of disinflation is lessened, as the need for the monetary authorities to show their resolve by accentuating their restrictive stance[15] is reduced. In the real world, wage indexation operates with lags, increasing the backwardness of wage and price adjustment; therefore, the speed of disinflation could slow down.[16] In particular, the shift from an indexation mechanism based on past price changes to a framework where inflation targets play the pivotal role may lower the costs of disinflation.

The empirical analysis presented in the following sections is an attempt to provide an evaluation of the effects stemming from the incomes policy episode. This evaluation must be considered as preliminary because some crucial features of the new bargaining system are still unsettled and quite long lags characterise the adjustment to institutional changes. Furthermore, there is no analysis of the linkages between the incomes policy episode and the overall process of fiscal consolidation and social policy reform. On this issue, one argument sees the wage moderation and the fiscal adjustment so far implemented as strictly complementary, in that unions' cooperation provided the necessary consensus for implementing painful policies which could, otherwise, have resulted in costly social conflicts. An opposite argument is that the *concertazione* has rather been an obstacle to the fiscal adjustment process, because structural reforms have been watered down by the need to obtain unions' agreement, particularly in the social policy field (social security and labour market regulation). Indeed, to evaluate these aspects, the interplay between the behaviour and preferences of each social partner should be analysed in depth. These issues are well beyond the scope of this paper.

3. An empirical evaluation of the changes in aggregate wage behaviour

Many observers have deemed the reform in the wage bargaining mechanism as the main factor lying behind the overall decline in inflation since 1992, despite two episodes of sharp depreciation of the exchange rate.

---

[15] For a more complete discussion on the interactions between monetary policy credibility and pricing behaviour, see Ball (1995).

[16] See, for instance, Chanda et al. (1992).

Two types of evidence have been used so far to provide empirical tests of these hypotheses: careful evaluations of the pay rises established in the most important contracts agreed upon; econometric experiments looking at the stability of wage and price behaviour in the recent period.

Examples of the former are in ISCO (1997) and IRS (1997), where the rise in minimum contractual wages is compared to target inflation, which according to the July 1993 agreements had to act as a yardstick. While providing useful bits of information, this kind of exercise cannot identify the specific role of the new bargaining mechanism in shaping actual wage behaviour. In the long run real wage growth has to be related to productivity growth and the reference to target inflation has to be interpreted as a coordination mechanism, not as an automatic rule shaping the evolution of nominal wages.

The econometric experiments have mostly analysed the inflation performance which followed the 1992 ERM crisis. After the sharp devaluation of the lira, a substantial increase in inflation was commonly considered as a very likely outcome, given a weight of imports in domestic demand close to 0.2. The fact that, on the contrary, inflation slightly decreased, fluctuating around 4% in 1993, has been interpreted as evidence of a structural break in wage and price behaviour. The incomes policy agreements have been considered the main determinant of the better-than-expected inflation outcome, both directly as a source of wage moderation and indirectly through their possible effect on inflation expectations and price behaviour.[17]

Recently, several analyses based on the BIQM have cast serious doubts on the purport of the July 1993 agreement on the recent inflation. Two additional factors have been emphasised: the sharp fall in aggregate demand, brought about by the recession in Europe and the fiscal squeeze; the delay in the pass-through of the Lira devaluation to import prices.[18] The former, in particular, can be considered as an alternative source of wage moderation.

Our experiment uses the BIQM analytical framework but differs from the aforementioned works in three respects. First, it takes advantage of the longer span of data now available. Secondly, it tries explicitly to identify the appearance of a new regime in wage behaviour, instead of relying uniquely on tests for generic structural breaks in the wage equation. It considers both hints of changes in the NAIRU and changes in the dynamic adjustment of prices to wages. Finally, it characterises the effects of the incomes policy agreements as a temporary deviation of wage behaviour from its long-run pattern that seems significant on economic grounds. This evaluation of the incomes policy effects is then used for simulations of the whole BIQM, the overall structure of which is described in the Appendix; such simulations allow us to disentangle the effects of the incomes policy from those due to other exogenous shocks, taking into account all possible feedbacks.

The wage equation we use is a Phillips curve,[19] describing a disequilibrium adjustment process in the labour market: wage inflation depends on price inflation as well as on the rate of unemployment. Price inflation is measured both by expected (as measured by survey results) and realised inflation, with a long-run homogeneity restriction. The use of an explicit measure of expected inflation introduces some structure into the wage-to-price adjustment process instead of relying on a

---

17 The first channel is obvious. Concerning the second, Gavosto et al. (1994) suggest that firms, knowing that nominal wages of competitors were not reacting to the exchange rate devaluation, were more cautious in transmitting the exchange rate shocks to domestic prices. Moreover, it has been claimed that the dismissal of indexation made consumers more choosy. In a longer-term perspective, Pecchi and Piga (1997) argue that the absence of indexation widens the constituency of price stability supporters and strengthens the anti-inflationary resolve of economic policy.

18 Gavosto et al. (1995), analysing the BIQM, argue that there is not clear-cut evidence of structural breaks, as far as the pricing behaviour and the determination of wages and inflation expectations are concerned. Furthermore, Nicoletti-Altimari (1995), studying Italian survey data on expectations, finds strong evidence against the hypothesis of a change in the mechanism of expectation formation.

19 On the choice between the Phillips curve and the real wage equation, see Bean (1994) and Blanchard and Katz (1997).

distributed lag specification which would blur the distinction between the formation of inflationary expectations and the lagged responses to actual inflation;[20] unfortunately, our measure of expected inflation refers to a short-run horizon. Labour market tightness is measured by the rate of unemployment and by the degree of capacity utilisation, acting as a proxy for vacancies. An indirect measure of union power is introduced via the number of working hours lost due to strikes. As any standard Phillips curve, the equation does not allow for a direct response of wages to productivity growth.[21]

The specification includes also some additional factors which might lead to changes over time in the NAIRU:[22] a proxy for the replacement ratio and the ratio of North-South unemployment rates. The former should increase wage pressures for a given level of unemployment;[23] the latter takes into account the claim that unemployment in the Northern regions is the main factor affecting the bargaining process (hence the variable should bear a negative sign).[24] After some experiments with several lags for the national unemployment rate, the final specification includes a 5 quarters moving average.[25]

The adjustment of wages to prices has been refined considering the possibility of asymmetrical responses to positive and negative inflationary surprises. The initial specification was one where wages react to past inflation and the inflation expected for the current quarter, with a further catch-up term in case of divergence between expected and actual inflation. Some preliminary experiments showed an asymmetry in the effect of the catch-up variable, which was significant only in the case of positive unexpected inflation. This asymmetrical behaviour has been included in the final specification presented below.

The OLS estimates of the chosen specification are shown in Table 1. The results refer to a sample period ending in 1991, which is well before the incomes policy agreements.

The statistical properties of the model are satisfactory. The goodness of fit is surprisingly high, given that the estimated equation is in first differences and that the sample period is commonly thought to cover different bargaining regimes: the adjusted R squared is above 0.8 and the standard error of the regression is consistently below 1%. The LM test for autocorrelation confirms that this very simple specification captures quite well the dynamic properties of the endogenous variable, while the RESET test shows the validity of the linear specification. No sign of conditional or more general form of heteroskedasticity is present, suggesting that the OLS estimator is the best within the class of linear estimators, while the results of the Bera-Jarque test for normality ensure that the OLS estimator is not only robust but also efficient (conditional on the correctness of the exogeneity assumptions).

---

[20] In the BIQM, expectations are endogenised in a separate equation as a partly adaptive process with several additional variables taking into account cyclical and policy factors (see Visco (1984) and Nicoletti-Altimari (1995)).

[21] Hence, it implies that a productivity slowdown (acceleration) would lead to higher (lower) price inflation, for a given wage increase, requiring higher (lower) unemployment to reconcile workers' and employers' claims.

[22] We have chosen these variables on the basis of preliminary results emerging from a parallel research on the structural determinants of the NAIRU, currently under way. However, the results on the stability of the BIQM wage equation do not depend on the introduction of these additional variables.

[23] The replacement ratio is computed as the ratio between per-capita expenditure for unemployment benefits (including the benefits paid by the Wage Supplementation Fund and taking into account that only a minority of unemployed workers are covered by benefits schemes) and average per-capita earnings.

[24] In the specification search, we have examined several different measures of the North-South unemployment gap, always obtaining similar results. In the end, since our wage equation is embedded in a nation-wide econometric model, we have chosen a specification which includes both regional unemployment rates.

[25] This specification and the alternative one with only one term of unemployment are mostly equivalent. The former has been preferred because it produces a more regular wage growth pattern.

Table 1

Phillips curve

1973Q3-1991Q4

$$\Delta w_t = -0.441 - 0.0027\sum_{j=1}^{4} u_{t-j} + \pi_{t-1} + 0.342\left(\pi_t^e - \pi_{t-1}\right) + 0.160\left[\left(\pi_{t-2} - \pi_{t-2}^e\right) + \left|\pi_{t-2} - \pi_{t-2}^e\right|\right]$$

(-3.017) (-3.296) (2.495) (-1.993)

$$+0.111\Delta str_t + 0.468\left(\frac{CPU_t}{CPU_t + CPU_{t-1}}\right) + 0.247\left(\Delta RR_{t-1} + \Delta RR_{t-2}\right) - 0.275\sum_{j=2}^{6}\left(\Delta\frac{U_{t-j}^{CN}}{U_{t-j}^{S}}\right) + dummies$$

(2.378) (3.171) (1.533) (-2.250)

$\overline{R}^2 = 0.845$		$\sigma_\varepsilon = 0.0071$			DW = 2.40
Restriction test	:	F(1,61)	=	3.54	[0.065]
Autocorrelation (1 - 4)	:	F(4,58)	=	0.794	[0.534]
Heteroskedasticity (linear)	:	$\chi^2(11)$	=	8.052	[0.708]
Heteroskedasticity (expon.)	:	$\chi^2(11)$	=	17.54	[0.093]
Normality	:	χ^2 (3)	=	1.601	[0.659]
Functional form	:	F(2,60)	=	1.751	[0.182]
Chow test	:	F(20,82)	=	2.008	[0.015]
Harvey's	:	*t*(61)	=	0.989	[0.327]

Note: The variables are defined as follows: *w*, *u* and *str* represent the logarithm of the wage in the non-farm non-energy business sector, total unemployment rate and the number of working hours lost due to strikes, respectively; π denotes the inflation rate, π^e its expected value and (π^e-π) the discrepancy between actual and expected inflation, or the catch-up term; *CPU* is the index of utilised capacity; *RR* the proxy for the benefit replacement ratio; U^{CN} refers to the unemployment rate in the Northern and Central regions and U^S to the one in the South.

All the variables introduced in addition to the unemployment and the price inflation terms are significant (marginally for the replacement ratio) only in differenced form. Hence they do not affect the estimate of the long-run NAIRU. However, their effect on wage pressures is not irrelevant. In particular, the rise in the North-South unemployment ratio plays a significant role in the second half of the 1980s by offsetting the downward pressure on wages of the ongoing rise in overall unemployment.

The long-run NAIRU implicit in the equation[26] is 7.7, as against an actual unemployment rate averaging 11.5 since 1991. Taken at its face value, the persistent and widening difference between actual unemployment and the NAIRU suggests either the presence of a restrictive stance of economic policy over a very long period or the existence of extremely long lags in the process of adjustment of unemployment to its "equilibrium" level. It is not unlikely that both hypotheses bear some element of

[26] The NAIRU has been computed assuming a quarterly rate of growth of productivity equal to 0.5% and applying the following formula:

$$\text{NAIRU} = \exp\left(\left(\text{constants} - \text{productivity growth}\right)/\text{real rigidity}\right)$$

where the term constants represent the sum of the intercept and of the coefficient of the relative degree of utilised capacity, which in steady state collapses to a constant equal to one, and real rigidity is the inverse of the coefficient of the unemployment rate. Note that since a unique value for productivity growth is used, a productivity slowdown that implies the NAIRU would be overestimated in the early part of the sample and underestimated in the later part.

truth. Since 1989, average output growth has remained approximately one percentage point per year lower than the OECD average (and also lower than in Continental Europe): this outcome could be traced back to the large fiscal adjustment packages introduced almost yearly and to the restrictive stance of monetary policy maintained throughout the period. In turn, there is plenty of evidence confirming the very slow adjustment in the labour market; a quite simple example is provided by the very small and delayed (with respect to the output cycle) reduction in unemployment experienced at the end of the 1980s.

Besides the usual caveat concerning the statistical precision of NAIRU, one has to keep in mind the already mentioned limitations of the Phillips-curve specification in identifying structural shifts in the NAIRU.

As far as the adjustment to prices is concerned, nominal wages react quite quickly, with an elasticity of about 0.7 with respect to one quarter lagged inflation, a value which is not far from the coverage with respect to increases in the cost of living provided, on average, by the automatic indexation mechanism; the elasticity with respect to expected inflation is about 0.3 and the homogeneity restriction is not rejected by the data. As already said, the catch-up mechanism implies that wages recover the purchasing power losses due to past unexpected inflation, while there is no effect when negative price surprises arise. This is a possible source of inflationary bias; however, in quantitative terms, the effect seems quite small (wages recover in two quarters only one third of their purchasing power losses).

The following step is to test whether the incomes policy agreements and the reshaping of the bargaining framework have significantly modified wage behaviour. To verify this hypothesis, a comprehensive check of the stability of the Phillips curve has been carried out, considering several pieces of evidence potentially able to detect different types of structural shifts in the wage equation.

As the timing of structural breaks cannot be established a priori[27] and the analysis of the stability of the wage equation over a longer time span is by itself of interest, it is convenient to start from a test aimed at detecting instability of a general form. The test performed here (suggested in Hansen (1992)), is approximately a Lagrange multiplier test of the null of constant parameters against the alternative that the parameters follow a martingale.[28] It has the advantage of being very simple to compute and of detecting the parameters responsible for the break, but it is not designed for determining the timing of the change if one has occurred. Table 2 presents the results of this test applied to the Phillips curve estimated over a sample period extended to 1996Q4, both for the most relevant coefficients and for the equation as a whole. By comparing each statistic with the critical values, this test does not support the view that the incomes policy episode and the 1993 reform represent fundamental innovations in the wage formation mechanism. None of the relevant parameters shows evidence of systematic changes and the equation itself seems stable.

As a further check, recursive estimates of the most relevant parameters of the Phillips curve were carried out[29] extending the sample to 1996, focusing on both the most important parameters and the implicit level of the NAIRU (Figure 3). There is evidence of a rather pronounced decline in the absolute value of the coefficient of the unemployment rate over the period 1982-86 and of a moderate rise afterwards, which seems more pronounced in the most recent period.

---

27 The implementation of the incomes policy and the reform of the bargaining institutions are distributed over a wide span of time: the dismissal of the automatic indexation mechanism dates back to December 1991; the first round of wage contracts set according to the new bargaining framework took place in 1994; firm-level bargaining peaked in 1995-96 and the issue of the discrepancy between target and actual inflation produced a tough confrontation in the second half of 1996.

28 The alternative incorporates simple structural breaks of unknown timing as well as random walk parameters. The more traditional CUSUM test was not deemed appropriate because, even though it aims at revealing instability of a general form, it is essentially a test to detect time variation in the intercept.

29 The starting estimation covers a nine year period; the recursive estimates therefore start from 1983.

Notwithstanding some wiggling around in the constant term, most of these changes in the unemployment coefficient are mirrored by fluctuations in the NAIRU estimates, which rises by 1.5 percentage points (to about 7.5) during the 1982-86 period, fluctuates around that level until 1992 (while actual unemployment was trending upwards), falling to around 7.0 since then (with a further widening of the difference between the actual unemployment rate and its "natural" level).

Table 2

Hansen's stability tests

Constant	0.037
Unemployment rate	0.035
Lagged inflation rate	0.049
Expected inflation rate	0.082
Catch-up	0.034
Degree of utilised capacity	0.038
Stability of the regression	1.989

Note: The 5% critical value for the Hansen procedure is 0.470 for tests on the stability of a single coefficient and 2.96 for a regression with 12 explanatory variables.

While there seems to be a structural improvement in the labour market following the incomes policy episode, the high imprecision of this result has to be stressed. By applying the delta method,[30] it turns out that the 95% confidence interval around the point estimate of the NAIRU is slightly more than 2 percentage points. Hence, the fluctuations of the estimated variable are not large enough to reject the hypothesis of a constant NAIRU.

The parameters capturing the adjustment of wages to prices appear to be somewhat more variable over time. The elasticity of wages with respect to the one-quarter lagged price change declines from a high of 0.85 at the beginning of the 1980s to just above 0.6 ten years later and this evolution resembles the changes in the degree of indexation provided by the *scala mobile.* By the same token, the rather large swings of this parameter emerging in recent years could be the counterpart of the abolition of the *scala mobile* clause.

The most interesting results are provided by the Chow test for post-sample goodness of fit, applied to the 1992Q1-1996Q4 period. The analysis of the t-statistic for each one-step-ahead forecast error shows that only in one quarter out of 20 (namely 1995Q2) the error is significantly different from zero. Moreover, a closer look at the data indicates that the wage change for 1995Q2 is definitely an outlier (equal to -0.4%, representing the only in-sample decline in the nominal wage), due to the large *una tantum* payments awarded in several sectors at the very beginning of 1995; in fact, on purely statistical grounds, the low P-value of the Chow test can mostly be attributed to this outlier.

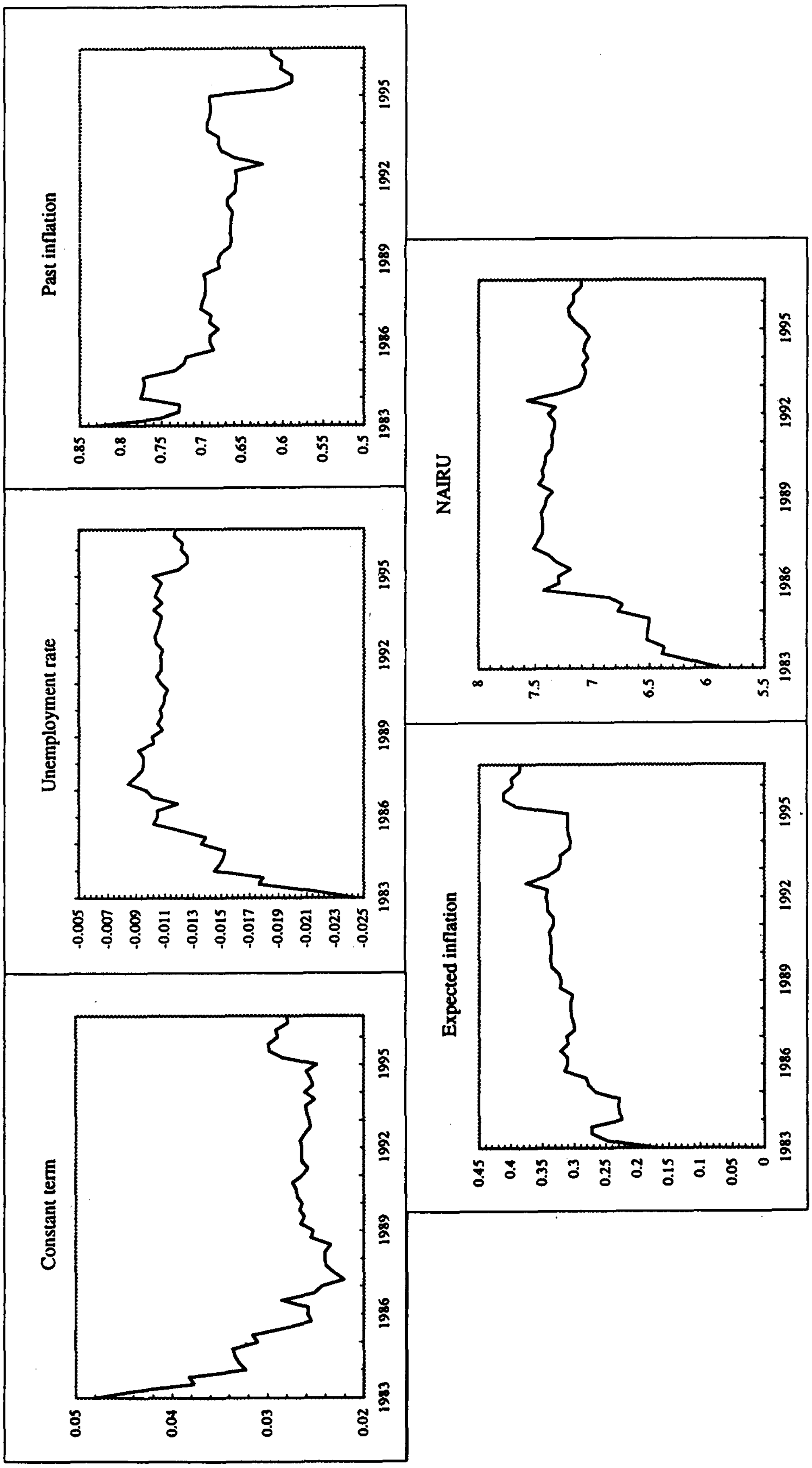
30 The delta method for computing the standard errors of a function of estimated parameters can be simply summarised as follows. Let $F = F(\theta)$ be an r x 1 vector function of the parameters θ of the model, and assume it is first-order differentiable. The estimate of F and its corresponding variance matrix are computed according to the formulae:

$$\hat{F} = F(\theta) \text{ and } \hat{V}(\hat{F}) = \hat{\sigma}^2 [\partial F(\theta)/\partial\theta]_{\theta=\hat{\theta}} \hat{V}(\hat{\theta}) [\partial F(\theta)/\partial\theta]'_{\theta=\hat{\theta}}$$

where $\hat{\theta}$ represents the maximum likelihood estimate of θ and $\hat{V}(\hat{\theta})$ a consistent estimator of its variance matrix. See also Rao (1973).

Figure 3

Recursive coefficients of the modified BIQM Phillips curve



However, it is possible to argue that errors which are not statistically significant may be significant in economic terms, given the lower inflation that has characterised the recent period: a 1% standard error in nominal wage growth is much more important in the recent period, with nominal wages growing at 3-4% per year, than in the 1980s, when they were growing at an average of 10%. Moreover, it has to be stressed that the forecasting errors show a rather systematic pattern: they are negligible in 1992 and, to a lesser extent, in 1993, become significantly negative during the 1994-95 period (actually peaking in 1995), and then reverse to positive values in 1996. Therefore, the equation estimated up to 1991Q4 appears to overestimate the wage evolution in 1993-95 and to underestimate it in 1996. According to this evidence the incomes policy would seem to have led to a temporary wage moderation, with a subsequent recovery of the real wage growth. This pattern fits quite well the standard picture of the incomes policy experiences since the 1960s.[31]

All in all, the empirical evidence gathered through the stability analysis of the Phillips curve can be summarised in three main findings. First, there is some weak evidence of a reduction in the NAIRU. However, because of the statistical imprecision of this result and the short period of time elapsed since the new bargaining system has been established, it is necessary to wait for additional evidence.

Secondly, there is stronger evidence hinting at quite relevant changes in the adjustment of wages to prices. This aspect, presumably related to the abolition of the old indexation mechanism and the new timing of the national-level contracts, will be further analysed in the next section.

Finally, there is a systematic pattern in the one-step-ahead wage forecasts errors, with an overestimation up to 1995 and a subsequent under-estimation. This characterisation suggests that, on top of any possible long-run changes in wage behaviour, the incomes policy episode may have played an autonomous role in the disinflation process. The fifth section will be devoted to measuring its effects.

4. A regime change in the adjustment of wages to prices?

In Section 2 it was argued that the absence of indexation and the two-year horizon of the national contracts could have increased the inertia of nominal wages (in levels), slowing down the nominal adjustment and raising the weight of expected inflation in the determination of nominal wages. The reduction of the short-run impact of inflationary shocks could have been associated with a longer feed back, producing a longer memory in the inflation process.

The empirical evidence from the stability analysis presented above somehow confirms this hypothesis. We now attempt to model explicitly the appearance of a new regime in the nominal wage adjustment, focusing on two aspects: the relative importance of expected and actual inflation and the quantitative size and length of the catch-up process.

A preliminary issue to be considered is that expected inflation (as actually measured by survey results) and target inflation (as fixed by the Government) may play different roles. In the long run, any sensible forward looking wage equation should include the expected inflation rate as the variable driving employers' and unions' behaviour. The targets fixed by the Government should play a role only insofar as they influence the private sector's expectations. A complete model should look at the interactions between target and expected inflation in affecting wages, but the brief period of time so far elapsed prevents an evaluation of their specific roles. In the short period considered here, the explicit reference to the inflation targets spelled out in the incomes policy agreements and the actual developments in wage bargaining suggest that the inflation targets may have been the reference variable for industry-level negotiations. Hence, the equation estimated in this section includes target

31 See Ulman and Flanagan (1971).

inflation, with a slightly better overall fit; the results obtained, however, are qualitatively robust to the alternative use of any of the two variables.

The following two-regime specification has been utilised:

$$\Delta w = (1 - a)\pi_{-1} + a\pi^e + b\left\{(\pi - \pi^e)_{t-i} + \left|(\pi - \pi^e)_{t-i}\right|\right\} / 2 + f(x) \tag{1}$$

$$\Delta w = (1 - a')\pi_{-1} + a'\pi^T + b'\left\{\sum\left[(\pi - \pi^T)_{t-i} + \left|(\pi - \pi^T)_{t-i}\right|\right]\right\} / 2 + f(x) \tag{2}$$

The first equation, already presented in the previous section, represents the behaviour prevailing in the 1970s and the 1980s, while the second picks up the new regime. Apart from the use of target inflation[32] in the second regime and expected inflation in the first regime, the two specifications may differ in three aspects: the size of a and a'; the size of b and b'; the length of the catch-up process, identified through a standard specification search. Having chosen to focus on the features of nominal adjustment, the specification is such that it keeps the coefficient of the variables included in the $f(x)$ term (described in the previous section) constant over the whole sample period.

The OLS estimates are shown in Table 3. The sample period extends from 1973Q3 to 1996Q4. Given the results of the stability analysis presented above and the ongoing nature of the changes in the bargaining system (with a first agreement to dismantle the *scala mobile* in July 1992 and a final agreement twelve months later), the break between the two regimes has been set in the first quarter of 1993. This choice allows a four-year period for testing hypotheses about the second regime.

The overall performance of the equation is quite satisfactory. The statistical fit, whether measured by the adjusted R-square or by the standard error of the regression, is good; there are no signs of heteroskedasticity, functional form misspecification or non-normality; all in all, no sign of misspecification appears, but for some evidence of first-order autocorrelation. As far as the coefficients of unemployment and the other real variables are concerned, the results are very close to the ones already shown in Table 1. In contrast, the difference between the two regimes appears rather extreme as far as the nominal variables are concerned.

In the second regime, nominal wage changes seem to be shaped only by the inflation target.[33] There is a stronger and longer catch-up to past inflation surprises: the adjustment is completed in 18 months and the size of the estimated coefficient (around 1.4) shows an overcompensation of past positive differences between actual and target inflation. On qualitative grounds, these characteristics correspond quite well to the priors of a more forward looking behaviour in the short run, with longer-lasting effects of inflationary shocks. However, the point estimates look quite implausible. The absence of any short-run role of actual inflation might simply be a temporary feature, due to the emphasis on target inflation as a coordination device during an economic turmoil period (target inflation was deemed to define the "fair" wage growth)[34]. The size of the catch-up effect might be picking up the normal surge in wage growth following a period of politically induced restraint.

More generally, the two caveats already mentioned have to be borne in mind: the absence of an explicit wage-to-wage spillover effect in the specification and, more important, the paucity of data allowing only a very preliminary evaluation.

---

32 On an annual basis, target inflation has been computed as a simple average of the values set by the Government one and two periods in advance. The quarterly series is obtained by using expected inflation as an indicator.

33 The unit restriction on the coefficient of target inflation is not rejected by the data.

34 As already said, the role of target inflation per se is uncertain in the long run.

Table 3

Two-regime Phillips curve
1973Q3-1996Q4

$$\Delta w_t = -0.338 - 0.00238 \sum_{j=1}^{4} u_{t-j} + \pi_{t-1} DUBF931 + 0.323\left(\pi_t^e - \pi_{t-1}\right) DUBF931$$
(-2.728) (-3.010) (2.373)

$$+0.170\left[\left(\pi_{t-2} - \pi_{t-2}^e\right) + \left|\pi_{t-2} - \pi_{t-2}^e\right|\right] DUBF931 + \pi_t^T (1 - DUBF931)$$
(2.050)

$$+0.717 \sum_{j=2}^{7}\left[\left(\pi_{t-j} - \pi_{t-j}^T\right) + \left|\pi_{t-j} - \pi_{t-j}^T\right|\right](1 - DUBF931) + 0.009 \Delta str_t$$
(2.126) (2.182)

$$+0.411\left(\frac{CPU_t}{CPU_t + CPU_{t-1}}\right) + 0.284\left(\Delta RR_{t-1} + \Delta RR_{t-2}\right) - 0.255 \sum_{j=2}^{6}\left(\Delta \frac{U_{t-j}^{CN}}{U_{t-j}^{S}}\right) + dummies$$
(2.869) (1.779) (-2.263)

$\overline{R}^2 = 0.848$		$\sigma_\varepsilon = 0.0074$			DW = 2.46
Restrictions test	:	F(2,79)	=	1.926	[15.244]
Autocorrelation (1 - 4)	:	F(4,77)	=	1.371	[25.179]
Heteroskedasticity (linear)	:	$\chi^2(12)$	=	9.283	[67.858]
Heteroskedasticity (expon.)	:	$\chi^2(12)$	=	21.059	[14.859]
Normality	:	$\chi^2(3)$	=	1.340	[51.173]
Functional form	:	F(2 , 79)	=	1.491	[23.138]

Note: The variables are the ones already presented in Table 1. *DUBF*931 is a dummy variable equal to 1 up to the first quarter of 1993 and π^T is the inflation target.

5. The overall effects of wage moderation: some counter-factual scenarios

The analysis carried out so far has shown some interesting sign of parameter shifts that can be related to the reform of the wage bargaining framework. However, the evidence about both the possible reduction in the NAIRU and the appearance of a new regime for the adjustment of nominal variables does not seem robust enough to provide a sound basis for a quantitative assessment of their effects.

Actually, following Sims (1980, 1982) and Cooley et al. (1984a, b) one might argue that genuine regime changes are rare, because agents with rational expectations know the options available to policy makers and form probability distributions over the range of possible policy stances. Thus, rather than changes in parameters, one has particular draws from probability distributions that have already been integrated into the optimisation problems of individual agents.[35] In this section, an alternative approach which is more in line with the mentioned claims is, therefore, pursued. Using the evidence presented in Section 3 on the systematic pattern in the one-step-ahead forecasting errors of the wage equation, the incomes policy episode is characterised as a phase of wage moderation followed by a wage pick-up. The evaluation of the overall effects of this episode is then worked out resorting to simulations of the whole BIQM.

Various counterfactual experiments are performed to simulate the developments of the Italian economy as if the incomes policy episode had not occurred. The simulations extend from 1993 to 1997,[36] encompassing the period characterised by the effects of two sharp depreciations, the one related to the Autumn 1992 ERM crisis and the other suffered in the first half of 1995, mainly related to political instability.[37] The main goal of the analysis is to provide a quantitative assessment of the specific contribution of wage restraint due to incomes policies to counteract the large price shocks deriving from the exchange rate depreciations and to favour the disinflationary process since 1993.[38] Exogenous and policy variables, including monetary policy, are fixed at their historical values and the ceteris paribus assumption is extended to the residuals as well.[39] Only results since 1994 are presented, given both the lags that characterise the response of prices to wages and the small amount of wage moderation, as measured here, in 1993.

The component of wage restraint attributable to incomes policy is measured using the residuals of the Phillips curve. More specifically, the quantification is based on the assumption that the counterfactual absence of the incomes policy can be modelled by endogenising the Phillips curve estimated with pre-agreement data (i.e. the first regime of the previous section). According to this hypothesis, the one-step-ahead forecast errors should represent the amount of "wage moderation" engineered by the incomes policy agreement.

The experiment has been further refined using the weak evidence on the identification of a second regime in the wage equation to obtain a gauge of the "pure noise" component included in the forecast errors of the wage equation. The forecast errors of the second regime equation have been considered as an estimate of this component and treated as an add-factor in the simulations.[40] The overall procedure can be questioned because the estimates of the new regime are not entirely reliable, as already pointed out. However, the unreliability of the new regime refers more to its long-run

---

35 In Sims' words: "Permanent shifts in regime are by definition rare events. If they occurred often, they would not be permanent [...] Normally policy actions are generated by a mechanism that, from the point of view of the public, forms a more or less stable stochastic process [...]. The historical record is likely to be especially valuable in projecting the effects of such actions. Radically new types of policy are likely to be, and ought to be, discounted because it is recognised that their effects are uncertain". For a detailed treatment of the theoretical and empirical issues concerning the Lucas critique and its impact on economic analysis, see Hoover (1995).

36 Data for 1997 are based on Bank of Italy forecasts. Since we are not concerned with absolute values of forecasts but with comparative evaluations of counter-factual scenarios, this is unlikely to bias the results of the analysis.

37 Since the BIQM is a non-linear model, the results of the experiments depend on initial conditions. By simulating over a period of radical changes, it is in principle possible to obtain biased outcomes. However the effects of non-linearities are second order and can be safely ignored in most cases.

38 This means that we do not explicitly measure the effects of other aspects of economic policy. A survey of the 1980s experience on the interactions between the several facets of economic policy in the disinflation process can be found in Andersen (1994).

39 Since the above simulations are counter-factual experiments rather than forecasting exercises, residuals which constrain endogenous variables to their historical levels can be computed. Unless they contribute to design alternative scenarios, they are treated as exogenous variables and left unchanged.

40 This procedure has been utilised for the 1993-96 period while in 1997 the add-factor of the Phillips curve has been set to zero, given that the identification of a noise component would have been arbitrary.

properties than to its in-sample features, as shown by the fact that the residuals during 1993-96 resemble quite closely a white noise process.

The other equation used to identify the effect of the incomes policy episode is the equation for inflation expectations. The rationale for this choice is the possibility that the incomes policy episode has led to a change in certain features of the inflationary process, an hypothesis somehow supported by the evidence presented in the previous section.[41] However, given the absence of a formal structural break in the equation included in the BIQM for endogenising inflation expectations,[42] the intervention made here has been simply that of fixing at zero the add-factor of the equation, so as to neutralise any indirect effects of the incomes policy agreements. The third and final intervention has been on the ratio of public to private wages. Its steady decline since 1991 (with a large pick-up in 1996) has been considered as part of the incomes policy framework and therefore neutralised by holding that ratio constant at its average level for the 1986-92 period.

The first simulation, henceforth labelled "no wage moderation" scenario, aims at measuring the overall impact of the changes in wage setting, in terms of inflation, growth and public finances. In this scenario monetary policy is kept unchanged; i.e. it does not react to the inflationary impulses coming from wages. The definition adopted for an unchanged monetary policy is over-simplified, given that no exact feedback rule for monetary policy (with forward and/or backward elements) is introduced into the model. Given the actual features of monetary policy in the period following the Italian exit from the ERM in September 1992,[43] the real interest rate has been kept at its historical values, with the exchange rate reacting to the new scenario, taking on board the effects of the higher nominal interest rates.

One additional limitation of this procedure is the assumption of an exogenous risk premium on the domestic exchange rate, which is the component explaining most of the large swings in the exchange rate during the period under scrutiny: the lira exchange rate was very sensitive to news about public finance, with a worsening likely to impose a severe penalty on the default risk premium and thereby induce a weakening of the currency. Therefore, the results worked out in the "no wage moderation" scenario are likely to underestimate the actual impact of the wage moderation episode. This is because an unchanged stance of monetary policy (as measured by real interest rates), in the face of additional inflationary pressures arising from a stronger wage growth, could have been interpreted by financial markets as too relaxed. It could, then, have caused a further depreciation of the lira and fuelled inflation. To tackle this possibility a second counter-factual experiment, henceforth called "risk premium shock" scenario, has been run assuming that the increase in nominal interest rates is entirely transmitted to the risk premium and hence to the exchange rate.[44]

Finally, a third experiment, labelled as "controlled inflation" scenario, considers the effects of the monetary authorities reacting to the ongoing inflationary pressures. It fully translates into output losses the additional inflationary pressures induced by the absence of wage moderation. In the experiment the overnight real interest rate is used as the policy instrument for targeting the yearly inflation rate over the whole period.[45]

Table 4 shows the results, computed as differences with respect to the baseline simulation. In the first two scenarios, as expected, wage growth is higher than in the baseline. The

---

41 More generally it has been argued that the income policy episode was conducive to a strengthening of the culture of monetary stability, a change which should have been reflected in the process of expectations formation.

42 See Nicoletti-Altimari (1995).

43 For a description of continuity and changes in monetary policy rules since 1992, see Passacantando (1996) and Visco (1995).

44 Since a shock to the risk premium feeds back to inflation and nominal rates, its amount was computed by trial and error.

45 The time path of the instrument was chosen so as to generate a smooth path for changes in the nominal interest rate.

difference widens quickly at the beginning of the period, peaking in 1995; later on there is a deceleration, and at the end of the simulation horizon (1997) wage growth becomes lower than the actual one. As a consequence, inflation, after having reacted more than in the baseline to the depreciations of 1992 and 1994-95, declines slightly less; at the end of the simulation period the deviation from the baseline tends to disappear, confirming the idea that the effects of the incomes policy agreements were temporary. The higher inflation affects GDP growth negatively, notwithstanding a moderate rebound at the end of the simulation period.

Table 4

Simulation results

Differences with respect to the baseline simulation

	"No wage moderation" scenario (A)				**"Risk premium shocks" scenario (B)**				**"No wage moderation" scenario with targeted 1996-97 inflation (C)**			
	1994	**1995**	**1996**	**1997**	**1994**	**1995**	**1996**	**1997**	**1994**	**1995**	**1996**	**1997**
Consumption deflator[1]	0.1	1.3	1.8	-0.1	0.3	1.5	2.6	0.8	0.0	0.0	0.0	0.0
Private sector wage growth[1]	0.7	4.0	1.3	-2.2	0.7	3.8	1.9	-1.3	0.4	2.8	-0.2	-2.1
GDP[1]	0.2	-0.2	-0.9	0.3	0.2	-0.1	-1.0	0.1	-0.1	-1.9	-1.1	2.6
Domestic households consumption[1]	0.5	0.0	-0.7	0.1	0.5	-0.1	-1.1	-0.3	0.1	-1.6	-0.2	3.4
Total investments[1]	0.5	-0.2	-2.3	-0.8	0.5	-0.2	-2.5	-1.3	-0.7	-5.8	-5.5	2.9
Total employment[1]	0.1	0.1	0.2	-0.1	0.1	0.1	-0.2	-0.1	0.1	-0.1	-0.3	0.0
Competitiveness of exports[1]	-0.2	-0.5	-1.1	1.4	-0.1	0.0	0.0	1.2	-1.5	-1.2	0.0	3.2
Current account balance to GDP (%)[2]	-0.2	-0.3	-0.2	0.0	-0.2	-0.3	-0.2	0.1	-0.8	-0.3	0.7	0.6
Net indebtedness to GDP ratio (%)[2]	0.7	1.0	2.1	2.9	0.7	1.0	2.2	3.3	1.4	4.3	6.5	4.3
Public debt to GDP ratio (%)[2]	-0.2	-0.7	0.5	3.5	-0.3	-0.9	-0.3	2.5	0.8	6.7	13.5	14.7
Treasury bill rate[2]	0.1	0.9	2.1	0.6	0.1	1.1	2.8	1.7	3.8	6.4	1.7	-2.9
Lira/DM[3]	-0.1	-0.2	-0.4	2.3	0.1	1.0	1.7	3.0	-2.3	-2.8	-0.9	5.5

(A) Old regime, Phillips curve for the private sector wages; public sector wages linked to private ones; real interest rate equal to the baseline.

(B) As (A), but eliminating the effects of nominal interest rates on exchange rates.

(C) As (A), but with interest rates targeting annual inflation in 1996 and 1997.

[1] Differences between rates of change. [2] Differences between levels. [3] Percentage differences between levels.

Looking in more detail at the first scenario, one can see that the additional inflation, as measured by the domestic consumption deflator, is nearly 1 percentage point (yearly average, 3.1 percentage points cumulated), reaching a maximum of 1.8 in 1996. This result turns out to be the balance of two contrasting effects: on the one hand, a higher wage inflation and, on the other hand, a temporary squeeze of firms' profit margins. The exchange rate only starts to depreciate (by 2.3%) at the end of the simulation horizon[46] and its immediate effects are partly offset by a 1% contraction in the pass-through. The higher inflation path negatively affects the level of economic activity (which by

[46] As already said, the ability to endogenise the exchange rate is quite limited. This is confirmed by the fact that most of its actual variation during the period is unexplained by the equation and appears as a residual, the path of which has been kept equal in all simulations but the second one.

itself helps to moderate inflationary pressures): in the final year GDP is 0.6% lower than in the baseline. The most important channel is the capital losses on financial wealth caused by the higher inflation, which depress disposable income and consumption.[47] Because of the greater interest payments, due to higher nominal interest rates, and, above all, the larger wage bill of the public sector implied by the assumption on the public to private wage ratio, net indebtedness over GDP diverges more and more from the baseline. Yet, the general government debt to GDP ratio is mostly unchanged or even falls at the beginning, thanks to the higher nominal GDP; only in the last two years of the simulation range does it start to grow at a very fast pace, exceeding by more than 3 percentage points the baseline value.

In the second scenario, all these effects are magnified by the higher responsiveness of the exchange rate to inflationary pressures: inflation is on average 1.3 percentage points higher than in the baseline (peaking at 2.6 percentage points in 1996); cumulated output losses reach 1.2%. Concerning public finances, the end-of-simulation indebtedness to GDP ratio is more than 3 percentage points higher than in the baseline; the deterioration in the debt to GDP ratio is partly offset by the surge in inflation, which erodes the stock of debt. At the end of the simulation period inflation still remains 1 percentage point higher than in the baseline and a further devaluation is taking place. Therefore, convergence has still to be achieved, implying that even this scenario may somehow understate the overall costs of the absence of wage moderation.

The public finance deterioration is magnified when a more restrictive monetary policy stance substitutes for wage moderation in curbing inflation. In the first three years, every component of aggregate demand slows down: the increase in the cost of capital freezes the recovery of investment and the rise in real interest rates induces households to postpone planned expenditure, which more than offsets the positive wealth effect of the lower inflation; the exchange rate appreciates and worsens the price competitiveness of exports. The fall in GDP and disposable income feeds back into demand and reinforces the contractionary effect of monetary policy. Also because of the higher interest on public debt, the debt to GDP ratio is nearly 15 percentage points higher than in the baseline, an outcome which might have jeopardised the fiscal adjustment process, feeding back to undermine the credibility of monetary policy in obtaining those inflation targets and in reducing, in 1997, the short-term interest rate.

Bearing in mind the fact that there is no reason to believe that the final year of the simulations represents an equilibrium outcome in both the baseline and the alternative scenarios, our results suggest that the temporary wage moderation engineered by the incomes policy agreements contributed quite significantly to disinflation, at the same time reducing the output and employment losses otherwise required. The alternative of an even tougher monetary policy would have been more costly in the short run and, because of its negative impact on public finances, might have lacked full credibility. The incomes policy agreements may, therefore, have provided a further contribution by strengthening the credibility of the disinflationary process.

[47] No significant contribution comes from the small competitiveness loss (export prices grow less than consumer prices and their rise tends to be offset by movements in the exchange rate). Furthermore, no positive impulse comes from the temporary redistribution of income in favour of labour since it is not relevant in determining aggregate household consumption. The usual accelerator mechanism drives down investment, amplifying the contractionary effect.

Conclusions

The role of incomes policy – while almost neglected in the recent experience of other industrial countries – is a central issue in the current debate on the Italian economic situation. Advocates of incomes policy have seen it as a central pillar of a new "stability culture", whose appearance is considered as essential for Italy's participation in the forthcoming EMU. Critics have emphasised the deleterious aspects of the *concertazione*, which may have blurred the distinction between the institutional roles of the Government and the social partners and, more substantively, could have distorted the structural features of the fiscal adjustment process.

In this paper we have focused on the direct linkages between the incomes policy episode, inflation and unemployment, neglecting the more general issues about its interactions with the fiscal adjustment process. In particular, the empirical analysis has aimed at measuring the specific relevance of the wage moderation episode, while considering whether there is any evidence of long-run changes in the NAIRU and in the features of the inflationary process deriving from the new bargaining framework. Our conclusions seem more balanced than those currently prevailing in the Italian debate.

We have found that the incomes policy agreements were conducive to a period of wage moderation, with some real wage pick-up later on, a quite standard characterisation of an incomes policy episode. This wage behaviour was important in counteracting the inflationary impulses stemming from two successive sharp devaluations and contributed to reducing inflation to levels not experienced in Italy since the second half of the 1960s. Depending on the different assumptions about the responsiveness of the exchange rate to the upsurge in domestic inflation, the absence of incomes policy would have pushed up consumer inflation by 2-3 percentage points in 1996 (with a cumulated 3-5 percentage points increase in the price level at the end of our simulation exercise, in 1997). Trying to obtain the same disinflation path without wage moderation would have asked for a much more restrictive monetary policy, whose impact on nominal GDP growth and interest payments would have jeopardised the consolidation of public finances (in terms of both net indebtedness and public debt to GDP ratios), putting at risk the credibility of (and the actual chance to reach) those inflation targets.

This contribution has to be considered of utmost relevance given the inflationary record of the Italian economy and the risk of undergoing an exchange rate crisis, thus feeding back an inflationary spiral. However, the evidence on long-run structural changes is less clear cut.

On both theoretical and empirical grounds we have found only weak evidence supporting the hypothesis of a reduction in the NAIRU associated with the new bargaining system. The small decline implicit in the Phillips curve estimates appears hardly significant from a statistical point of view; moreover, it is quite difficult to link it to the new bargaining system, given the multitude of factors which affect the NAIRU and the several facets of the new bargaining system which are still unsettled. Theoretically, it has been argued that the new system might lead to a reduction in the NAIRU through three channels: an increase in coordination in bargaining, a more cooperative stance by unions (more ready to internalise the implications of their behaviour), and the dismantling of the distortive aspects of the old indexation mechanism. However, only the third of these effects may be taken for granted (although its empirical relevance is less obvious). The second effect may be questioned in so far as the cooperative stance of unions may be insufficient to produce the differentiation in the wage structure, particularly along regional lines, which would be necessary to counteract the labour market segmentation plaguing Italy. Concerning coordination, no trend towards a more synchronised pattern has yet emerged and the apparent strong coordination across bargaining units so far obtained through the common reference to target inflation may be a temporary feature determined by the economic and political "emergency" situation experienced in Italy.

Actually, an important issue still unsettled in the bargaining system is the relationship between industry and firm-level bargaining. A shift towards an exclusively profit-sharing nature of the latter seems unlikely: firms' bonuses are increasingly important but they usually grant substantial wage rises almost irrespective of actual profitability, with limited variations over time. Insofar as the industry level contracts set wage rises tightly linked to target inflation, unions would be partly left aside in the distribution of productivity gains in the large portion of firms where decentralised bargaining does not take place. This would gradually lead to far-reaching changes in the role and nature of Italian unions. In the opposite polar case where industry contracts would result in real wage increases in line with productivity gains, the room left for decentralised bargaining would remain quite limited and the usual inflationary bias due to the overlapping of two levels of bargaining could emerge. Whatever the advantages and disadvantages of national or firm-level negotiations acting as a pivot in wage setting, our assessment is that an equilibrium bargaining structure has still to emerge.[48]

Qualitatively more clear-cut are the changes in the adjustment of wages to prices. We have argued that the absence of indexation and the two years horizon of the national contracts are likely to increase the inertia of nominal wages (in level terms), amplifying the weight of expected inflation in their determination. For a given level of expected inflation, the short-run impact of inflationary shocks should be reduced. These will somehow feed back later on, when renegotiation occurs. We have found some empirical evidence of this effect, even if its exact quantification based on the identification of a new regime in the Phillips curve appears somehow implausible.

Some ambiguities exist concerning the respective roles of target and expected inflation. The fact that the targets have been playing an effective role in driving nominal wages since 1993 is not something which may be taken for granted in the long run. To be effective on this respect, the inflation targets set by the Government must be credible to the parties involved in the negotiation, without becoming a matter of bargaining. Hence, the targets have to be a policy goal based on plausible forecasts. The independent role of the monetary authorities in warranting price stability is likely to have become more important for unions locked in by contracts with a long horizon.

Moreover, both the pros and the cons of a rise in nominal rigidity have to be considered. The pros have been emphasised by the recent experience: as said, the partial anchoring of nominal variables caused by the slow response of nominal wages to inflationary shocks played a crucial role in the disinflation obtained despite the two sharp depreciations. Among the cons one has to consider the possible fluctuations in real wages unrelated to the fundamentals of the economy; the longer memory in the inflationary process (the short-run impact of inflationary shocks is reduced and postponed, leading to later pick-ups in inflation); and the risk of frictions when renegotiation occurs.

The possible solutions to the issue of nominal rigidity are linked to the choice between firm and national level bargaining. If industry-level contracts maintain their central role, an evolution towards shorter-term wage settlements could counteract the "excessive" nominal wage rigidity. Conversely, a shift to a more widespread use of firm-level contracts should confine centralised bargaining to act only as a "safety net" for minimum wage and working conditions, in order to prevent the inflationary bias (and bitter conflicts) arising from the overlapping of two bargaining levels.

48 See Sestito (1996).

Appendix: The BIQM model

General overview[49]

In the BIQM, rather traditionally, a number of adjustment processes govern the short-run dynamics of the economy, with long-run behaviour consistent with a neo-classical model with exogenous growth. In the absence of shocks, when all adjustment processes have taken place and expectations are fulfilled – that is, in the long run – the model describes a full employment economy,[50] in which all real variables grow at the same rate, equal to the sum of the rates of growth of population and productivity. Output, employment and the capital stock are consistent with an aggregate production function; all relative prices are constant. Inflation is constant as well, and equals the exogenous rate of growth of foreign prices, assuming no changes in the nominal exchange rate. Money is neutral, but not super-neutral, and the model is dynamically stable.

The theoretical structure underlying the steady-state is also a traditional one. The supply sector can be thought of as being composed of producers who operate in a monopolistically competitive market for their output and are price-takers in the market for production factors, each producer being endowed with the same (Cobb-Douglas) constant returns to scale technology. Along a steady-state growth path, firms decide, in each period, the cost-minimising factor mix. The level of domestic activity is then set to generate, given factor demands, a non-accelerating-inflation rate of unemployment. Life-cycle consumers choose the desired addition to the real stock of total wealth. The latter must be consistent with the demand for new capital by the firms, the demand for net (real) foreign assets by both firms and consumers and the (real) addition to the stock of government debt. Given the latter, relative prices (the real interest rate, the real exchange rate and the real wage rate) are determined so as to achieve the required consistency. As consumers do not anticipate, in the computation of their life-time resources, the need for the government to satisfy a long-run solvency condition, the stock of government debt is perceived to be part of total wealth, and Ricardian equivalence does not hold.

The intrinsic dynamics of the long-run equilibrium, deriving from wealth and capital accumulation, combines with the dynamics coming from the short-run adjustment processes. The most important ones reflect the putty-clay nature of capital, the stickiness of prices and wages, the possibility that expectations differ from realised values, and the corresponding revision of both plans and expectations. As to expectations, the BIQM, to a large extent, makes use of survey data on actual expectations, namely for price and exchange rates changes. This allows the estimation of the model to be carried out without any need for arbitrary assumptions on the expectation formation mechanism.

The wage-price block

In the BIQM, aggregate demand deflators are given by weighted averages of the value added and import deflators, the weights being different case by case. Import deflators are essentially exogenous, being calculated by converting foreign prices into lire. The equation for the deflator of imported manufactures allows for the possibility that foreign firms, in order to defend their market shares, may price to market: the pass-through is modelled as a positive function of the Italian cycle as opposed to that of the rest of the world and as a negative function of the lira effective exchange rate. Exchange rates are endogenised through an uncovered interest rate parity, with exogenous risk premium and expectations of exchange rate changes modelled as functions of PPP, interest rate differentials and past forecast errors. The private sector value added deflator is determined within an oligopolistic competition framework: firms apply a mark-up over unit labour costs, the mark-up being

49 A more detailed description is in Galli et al. (1990).

50 Full employment means a level of production and employment consistent with a constant rate of growth of prices.

a function of competitiveness on both foreign and domestic markets and of demand pressures, as captured by the degree of capacity utilisation. In addition, the short-run mark-up is influenced by changes in the deflator of imported inputs and foreign firms' pass-through, the effect being asymmetric in both cases. Wages are modelled according to a wage-leadership scheme, with the private sector being the leader and the public, energy and agricultural sectors the followers; the private sector's wage dynamics is determined via a Phillips curve.

The aggregate demand block

While in the long run the process of capital accumulation determines the level of economic activity, in the short run the nature of the model is basically Keynesian, with output being demand determined and with typical situations of excess supply both in the goods and in the labour markets. Consumption of non-durables and durables are considered separately; the latter is modelled as a function of non-durables consumption, used as a proxy for permanent income, relative price and interest rates; the former is modelled according to the life-cycle theory, with disposable income and wealth both playing a role. Disposable income is computed in such a way as to take the Hicksian correction into account, with inflation having a negative influence through the capital losses on total financial wealth. Investment decisions are specified separately for machinery, structures and residential construction. Investment in machinery is designed within a putty-clay scheme, with delivery lags and expectation lags determining the dynamic structure. Foreign trade is modelled very much according to the usual hypotheses, with competitiveness and absorption as explanatory variables. Capacity utilisation, used as a non-price rationing variable, is found to have a significant effect on both.

References

Andersen, P. S. (1994): "Disinflation experience in Finland compared with other OECD countries", in J. Åkerholm and A. Giovannini (eds.), *Exchange rates policies in the Nordic countries*, CEPR, London, pp. 173-227.

Ball, L. (1995): "Disinflation with imperfect credibility". *Journal of Monetary Economics*, No. 35(1), pp. 5-23.

Banca d'Italia (1986): *Economic Bulletin*, No. 2, February.

Barca, F. and M. Magnani (1989): *L'industria tra capitale e lavoro,* Il Mulino, Bologna.

Bean, C. (1994): "European Unemployment: A Survey". *Journal of Economic Literature*, No. 32, pp. 553-619.

Blanchard, O. and L. F. Katz (1997): "What do we know and do not know about the natural rate of unemployment". *Journal of Economic Perspectives*, pp. 51-72.

Bodo, G. and I. Visco (1987): "La disoccupazione in Italia: un'analisi con il modello econometrico della Banca d'Italia". *Contributi all'analisi economica*, No. 3, December.

Bodo, G. and P. Sestito (1991): *Le vie dello sviluppo,* Il Mulino, Bologna.

Bodo, G., G. Bruno, G. Parigi and S. Siviero (1994): "Diagnostic Checking of large-size Econometric Models: an Application to the Models of the Bank of Italy". *Economic Notes*.

Calmfors, L. and J. Driffil (1988): "Bargaining structure, corporatism and macroeconomic performance". *Economic Policy*, No. 6, pp. 14-61.

Calmfors, L. (1993): "Centralisation of Wage Bargaining and Economic Performance: A Survey". *OECD Economics Department Working Papers,* No. 131.

Casavola, P. (1994): "Occupazione e disoccupazione. E' cambiato il mercato del lavoro italiano?" *Mimeo*, Bank of Italy.

Casavola, P., A. Gavosto and P. Sestito (1995): "Salari e mercato locale del lavoro". *Lavoro e relazioni industriali,* No. 4.

Chanda, B., P. Masson and G. Meredith (1992): "Models of Inflation and the Costs of Disinflation". *IMF Staff Papers*, No. 39 (2), pp. 395-431.

Checchi, D. (1995): "La moderazione salariale negli anni '80 in Italia". *Lavoro e Relazioni Industriali*, No. 2.

Commission of the European Communities (CEC) (1993): "The economic and financial situation in Italy". *European Economy - Reports and Studies*, No. 1.

Cooley, T. F., S. F. LeRoy and N. Raymon (1984a):"Econometric policy Evaluation: note". *American Economic Review*, No.74(3), pp. 467-70.

Cooley, T. F., S. F. LeRoy and N. Raymon (1984b): "Modeling policy interventions". Unpublished typescript, University of California, Santa Barbara, CA, and University of Missouri, Columbia, MO.

Dell'Aringa, C. (1995): "L'accordo di luglio 1993 e gli effetti sul costo del lavoro: le prime verifiche". *Rivista Internazionale di Scienze Sociali*, No. 3, pp. 501-24.

Demekas, D. (1995): "Labour market institutions and flexibility in Italy". *Labour*, No. 9(1), pp. 3-43.

De Stefanis, S. (1995): "Aggiustamento dei salari, indicizzazione e inflazione: un'analisi empirica per l'industria in senso stretto, 1963-85". *Lavoro e relazioni industriali*, No. 3, pp. 3-58.

Elmeskov, J. (1993): "High and persistent unemployment: assessment of the problem and its causes". *OECD Economics Department Working Papers*, No. 132.

Faini (1995): "Stesso lavoro, diverso salario?", in F. Galimberti et al. (eds.), *Le nuove frontiere della politica economica*, Il Sole 24 Ore, Milano.

Galli, G., D. Terlizzese and I. Visco (1990): "Short and long-run properties of the Bank of Italy Quarterly Econometric Model". *Dynamic Modelling and Control of National Economies*, pp. 440-506.

Gavosto, A., R. Sabbatini and P. Sestito (1994): "Inflazione e ciclo economico: alcuni elementi di analisi per interpretare i recenti avvenimenti in Italia". *Moneta e Credito*, p. 188.

Gavosto, A., S. Siviero and D. Terlizzese (1995): "Output Gap and Inflation in Italy". *Mimeo*, Bank of Italy.

Giavazzi, F. and L. Spaventa (1989): "Italy, the real effects of inflation and disinflation". *Economic Policy*, No. 4(8), pp. 133-172.

Gray, J. A. (1976): "Wage indexation: a macroeconomic approach". *Journal of Monetary Economics*, April(2), pp. 221-235.

Hansen, B. E. (1994): "Testing for Parameter Instability in Linear Models", in N. R. Ericsson and J. S. Irons (eds.), *Testing Exogeneity*.

Hoover, K. D. (1995): *Macroeconometrics: Developments, Tensions, and Prospects*. Kluwer Academic Publishers.

IRS-CNEL (1997): *Rapporti sui salari.*

ISCO (1997): *Rapporto semestrale*, No. 67. Rome.

Layard, R., Nickell, S. and Jackman, R. (1991): *Unemployment: Macroeconomic performance and the labour market*. Oxford University Press, Oxford.

Nicoletti Altimari, S. (1995): "Uno studio sulle aspettative di inflazione in Italia: 1970-1995", in *Ricerche quantitative per la politica economica, 1995*. Banca d'Italia (forthcoming).

Nicoletti Altimari, S., R. Rinaldi, S. Siviero and D. Terlizzese (1994): "Monetary Policy and the Transmission Channels in the Bank of Italy's quarterly econometric model", in *Financial structure and the monetary policy transmission mechanism*, BIS, Basel.

OECD (1997): *Employment outlook 1997*. Paris (forthcoming).

Passacantando, F. (1996): "Building an institutional framework for monetary stability: the case of Italy (1989-1994)". *BNL Quarterly Review*, March.

Pecchi, L. and G. Piga (1997): "Indexation, inflation and central bank independence". *Mimeo.*

Rao, C. R. (1973): *Linear statistical inference and its applications*, Wiley and Sons, New York.

Regini, M. and Regalia I. (1996): "Italia anni '90: rinasce la concertazione". *IRES Discussion Papers*, No. 23.

Robotti, L. (1973): "Incidenza dell'indennità di contingenza nella dinamica salariale in Italia: 1951-70". *Rassegna di Statistiche del Lavoro,* Supplemento, pp. 7-21.

Sestito, P. (1994): "Determinazione del salario in Italia: una rassegna della letteratura macro", in C. Dell'Aringa (ed.), *La determinazione dei salari: teoria ed evidenza empirica*, Edizioni Scientifiche Italiane, Napoli.

Sestito, P. (1996): "Contrattazione aziendale e piccole imprese: quali spazi dopo gli accordi di Politica dei Redditi?" *Conference of the Associazione Italiana per lo Studio delle Relazioni Industriali.*

Sims, C. A. (1982): *Policy Analysis with Econometric Models*. Brookings Papers on Economic Activity, pp. 107-152.

Sims, C. A. (1980): "Macroeconomic and reality". *Econometrica*, No. 48, pp. 1-48.

Siviero, S. and D. Terlizzese (1995): "Crisi di cambio e innovazioni nei comportamenti: alla ricerca di breaks strutturali nel modello econometrico della Banca d'Italia". *Ricerche quantitative per la politica economica*, Banca d'Italia (forthcoming).

Staiger, D., J. H. Stock and M. W. Watson (1996): "How precise are estimates of the natural rate of unemployment". *NBER Working Paper Series*, No. 5477.

Ulman, L. and R. J. Flanagan (1971): *Wage restraint: a study of incomes policies in Western Europe*. University of California Press, Berkeley.

Visco, I. (1984): *Price expectations in rising inflation*, North Holland, Amsterdam.

Visco, I. (1994): *Caratteri strutturali dell'inflazione italiana (1986-1991)*, in C. Dell'Aringa (ed.), *Caratteri strutturali dell'inflazione italiana*, Il Mulino, Bologna.

Visco, I. (1995): "Inflation, inflation targeting and monetary policy: notes for discussion on the Italian experience", in L. Leiderman and L. E. D. Svensson (eds.), *Inflation targets*, CEPR.

Comments on: "NAIRU, incomes policy and inflation" by Silvia Fabiani, Alberto Locarno, Gian Paolo Oneto and Paolo Sestito

---

by Sean Craig

This paper examines the question of the extent to which the change in incomes policy in 1991 contributed to Italy's good inflation performance after the ERM crisis led to large lira devaluations. This change involved a shift from a backward looking indexation mechanism where wages were linked to past inflation to a forward looking mechanism where wage increases are tied to the governments forecast of inflation. The approach of the paper is to model this change in the indexation mechanism as a shift in the estimated coefficients of the Philips curve equation in the econometric model of the Bank of Italy (BOI). It does this in three steps:

1) it tests whether the existing Philips curve equation has remained stable during the period of structural change and, thus, can be used to perform out-of-sample simulations;

2) it estimates a version of the Philips curve in which the specification of the equation is modified to reflect the change in the indexation mechanism by using dummy variables to allow the estimated coefficient on inflation to change;

3) it uses simulations of the BOI quarterly model to calculate how much higher inflation would have been had the Philips curve not changed. In Essence, this exercise involves a comparison of the simulations of two version of the model: one with the old Philips curve and the other with the modified specification.

Turning first to the test for structural change, the authors cannot reject the hypothesis that the Philips curve equation was stable over the sample period. While this is an encouraging result, it should be noted that their test for structural stability has low power in their relatively small sample, as the authors acknowledge.

The modification of the Philips curve equation to incorporate the change in the indexation mechanism is an interesting innovation. Unfortunately, the estimated parameters intended to capture this effect are implausible, casting doubt on the usefulness of the results. As the authors note, this may be due in part to the small number of observations available since the change in indexation occurred. Another problem is that the authors do not test the restrictions implied by this modification. One simple but revealing test would be to test (using an F-test) the specification incorporating the restrictions against the one without them to see whether the restrictions are rejected.

The last issue addressed by the paper is the impact of this change in the indexation mechanism on the inflation process. The comparison of the simulation results for the model with and without the modified Philips curve shows a very large effect of the change in the indexation mechanism on inflation – in the absence of the structural change, inflation would have been almost 2 percentage points higher in 1996. One concern must be that the paper attributes too much of the substantial improvement in Italian inflation performance to the change in the indexation mechanism, Other important structural changes that occurred at the same time may have contributed. In particular, structural reforms made it easier for firms to reduce employment which, in turn, contributed to strong productivity growth. As a result, unit labour costs actually fell in 1994 and 1995, despite the acceleration in wage inflation that began in 1994. Clearly, this and other structural changes could have had an impact on the estimated coefficients.

In conclusion, the modification to the Philips curve equation is unlikely to be sufficient to be isolate the indexation channel from these other channels through which structural changes could influence the inflation process. In particular, the likely importance of the change in employment practices (and the resulting effect on productivity) suggest that it might be appropriate to allow for structural shifts in other equations in the BOI model in addition to the Philips curve. The authors have developed and interesting approach to analysing the inflationary impact of structural shifts in labour markets, but they need to extend it to other sectors of the BOI model before they can adequately explain the dramatic improvement in Italian inflation performance following the structural reforms.

A look at the US inflation puzzle[1]

Cara S. Lown and Robert W. Rich

Despite weaker-than-anticipated inflation the past year, Federal Reserve Bank (of San Francisco) President Parry stressed continued vigilance on the price front. "I have a question mark, and it leads me to recommend vigilance with regard to inflation, but I do have to note that things have turned out well" – contrary to his expectations and most economic models, he noted.

He acknowledged being puzzled over the reasons why inflation has remained below expectations, despite robust growth, tight labor markets and rising wages.

"You've either been lucky, in which case the old relationships will reassert themselves, or you've got a new regime underway. And I don't think we know enough at this point to know which of those two things is operative," which means extra caution on inflation.

[Dow Jones News Service: 7th January 1997]

Introduction

Since the current expansion began in 1991, the inflation rate has remained at or below 3%. But for well over three years, expectations have generally foretold of an inflation pickup: the Blue Chip Consensus forecast (Eggert, various issues) has been predicting such a pickup, and both the yield curve and commodity prices have at times portended a pickup as well.[2] For example, the run-up in commodity prices from November 1993 to April 1995 fueled speculation of higher inflation rates; these higher rates did not materialize. The lack of an inflation increase has led some individuals to conclude that there is an "inflation puzzle".

One possible explanation for this "puzzle" is that it reflects a fundamental shift in the dynamics of the inflation process. Changes in labor market behavior, increased international competition, and changes in the way monetary policy is conducted have been cited as support for this explanation. Alternatively, one could argue that forecasters and market makers have simply "missed the boat," creating the perception of an inflation puzzle that does not exist.

The purpose of this paper is to investigate the issue of an inflation puzzle and, in particular, to assess the merits of the preceding arguments concerning its existence. We explore this issue by estimating Phillips-curve models for price inflation in the core consumer price index (CPI) and wage inflation measured by compensation growth and evaluating their forecast performance. A central part of the analysis focuses on the behavior of the price-inflation and wage-inflation series over the current expansion. Accordingly, we conduct a variety of tests for instability in the Phillips-curve models and any evidence of changes in the estimated relationships over the post-1991 period.

Our findings indicate that price inflation over the current expansion has *not* been unusually low relative to its historic proximate determinants. In particular, the results suggest that our

1 The authors are grateful to J.S. Butler, Gabriele Galati, Steve Kamin, Jonathan McCarthy, Patricia Mosser, Richard Peach and Charles Steindel for helpful comments on an earlier version of this paper. They have also benefited from the suggestions of conference participants at the Bank for International Settlements. Beethika Khan provided expert research assistance. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

2 Goodfriend (1993) contains a discussion of other "inflation scare" episodes in his case study of the Federal Reserve's interest rate policy from 1979-92.

price-inflation Phillips-curve model – modified to include unit labor costs – fits the data quite well, with stability tests and out-of-sample forecasts providing little evidence of a change in the behavior of inflation over the past few years. However, the results also indicate that compensation growth was markedly lower than the corresponding forecasts from our wage-inflation Phillips-curve model during the period 1992-94. While the current behavior of compensation growth appears to be consistent with the levels of other related macroeconomic variables, this relatively weak increase in compensation growth was a major contributor to the low level of inflation observed over the current expansion.

Our study also builds upon previous Phillips-curve research which has postulated that the dynamics of price and wage changes depend on both the level and rate-of-change of key aggregate demand variables such as the output gap (the log ratio of actual to potential real GDP) or the unemployment gap (the difference between the actual rate of unemployment and the NAIRU – the nonaccelerating inflation rate of unemployment). For the price-inflation Phillips curve, the estimation results provide evidence of a statistically and economically significant rate-of-change effect for the output gap. In the case of the wage-inflation Phillips curve, we implicitly assume a constant NAIRU and observe that the estimated compensation growth equation performs quite well in terms of its within-sample predictions and out-of-sample forecasts. These findings suggest that the debate concerning time-variation in the NAIRU and a possible decline in its value during the 1990s may not be particularly important for gauging the behavior of compensation growth.

The outline of the paper is as follows. We begin by reviewing the recent behavior of inflation. We suggest reasons why forecasters might have been expecting an increase in the inflation rate and we also discuss factors that have likely helped to mute the inflation rate over the current expansion. In Section 2, we specify a price-inflation Phillips-curve model, present the estimation results and tests for stability of the equation over the post-1991 period.

We then continue our exploration into the issue of an inflation puzzle and examine whether the recent movements in inflation may instead reflect unusual behavior in its underlying determinants. In Section 3, we analyze the behavior of compensation growth in more detail. We specify and estimate a wage-inflation Phillips-curve model for compensation growth and document the weak increase in this variable that occurred from 1992-94. In Section 4, we investigate if any changes in the behavior of labor market variables can account for this previous shortfall in compensation growth. The final section concludes.

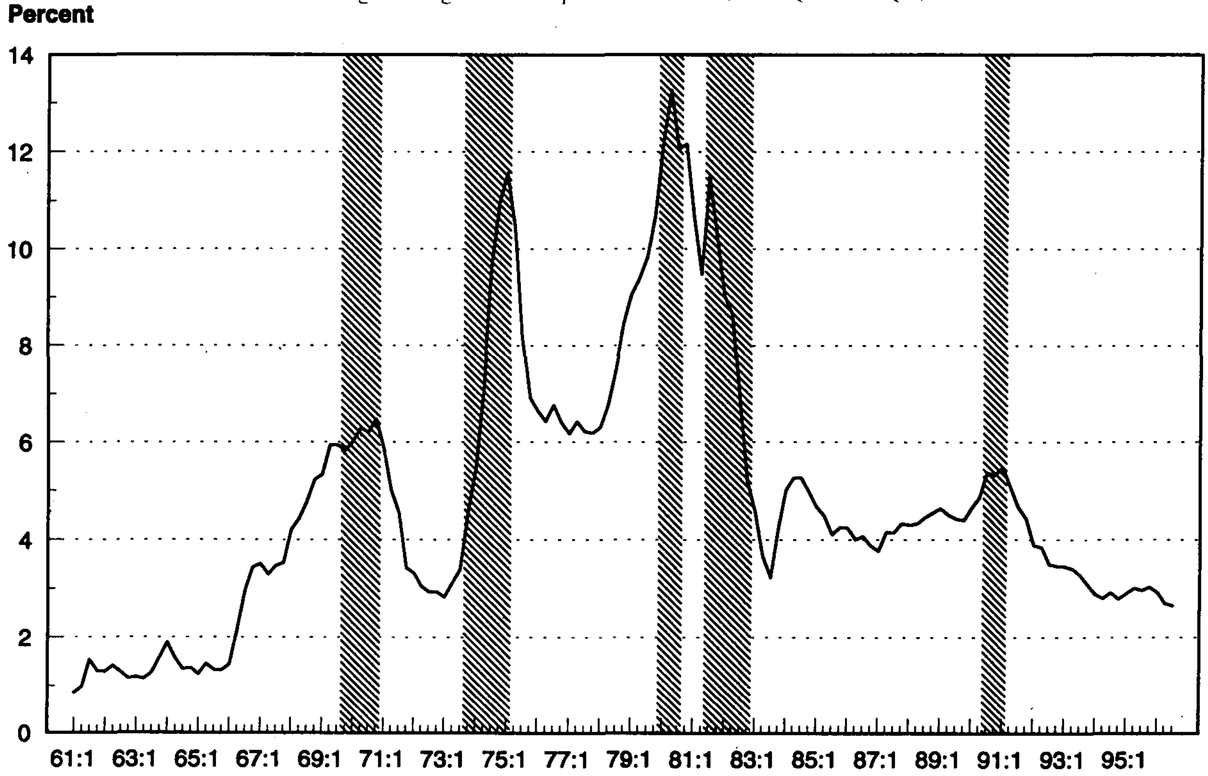
1. The nature of the "inflation puzzle"

Chart 1 presents the behavior of core CPI – the CPI excluding its food and energy components – since the early 1960s. As the chart shows, inflation has typically accelerated during each of the four previous expansions. In contrast, inflation was slowly declining during the first few years of the current expansion and, more recently, has remained roughly constant.

There are other reasons why the recent behavior of inflation might appear surprising. Several variables commonly regarded as inflation indicators have been at levels which historically have signaled an inflation pick-up. One such variable is the level of the actual unemployment rate relative to the NAIRU – the unemployment rate that is consistent with a constant rate of inflation. The civilian unemployment rate is shown in the upper panel of Chart 2, with an assumed value of 6% for the NAIRU. As the chart shows, the unemployment rate series has been below this threshold level since late 1994. Admittedly, there have been discussions about whether the NAIRU has declined somewhat during the 1990s. However, few researchers have suggested that the NAIRU has fallen to a level such that the unemployment rates observed since 1995 imply a *constant* rate of inflation.[3]

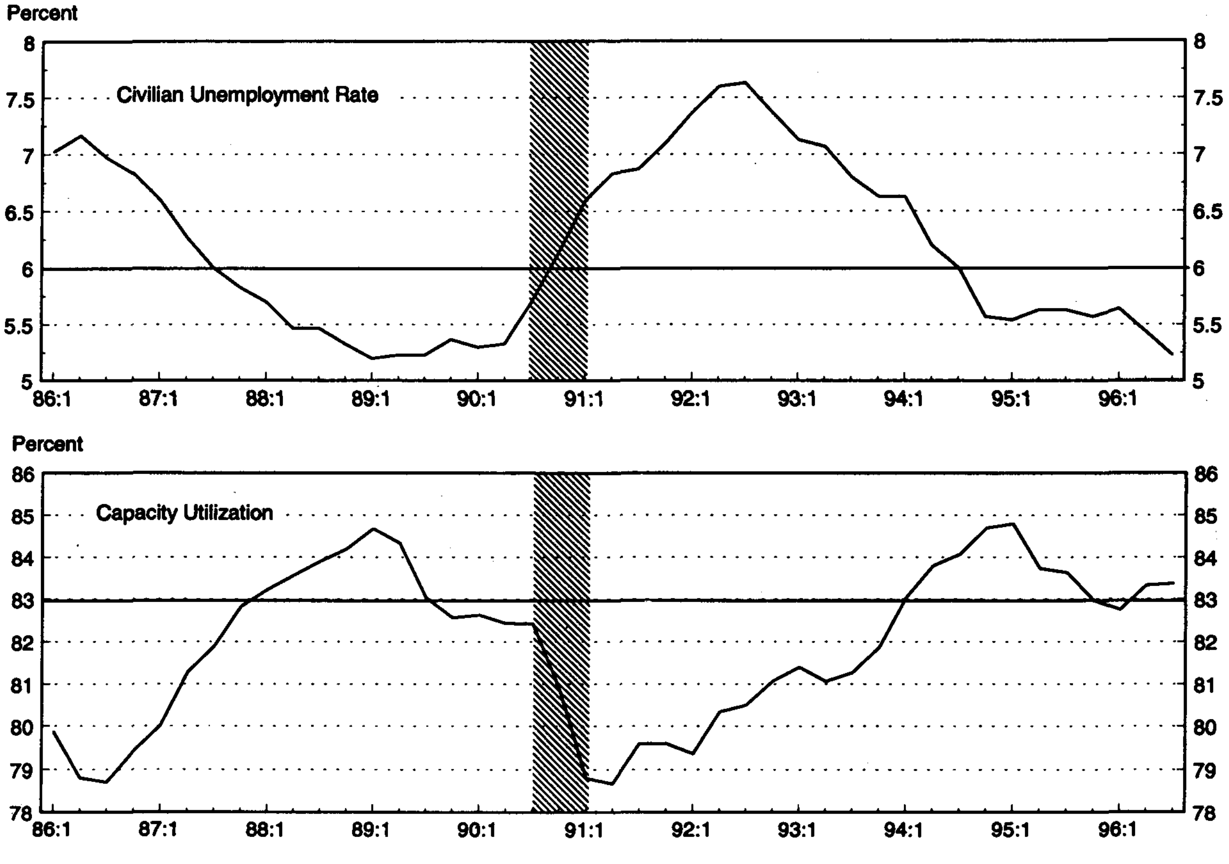
[3] One exception is Gordon (1996) who obtains an estimate of 5.3% for the NAIRU starting in 1996.

Chart 1
Core consumer price index
Percentage change from 4 quarters earlier (1961Q1-1996Q3)



Note: Shading refers to NBER recessions.

Chart 2
Inflation thresholds?

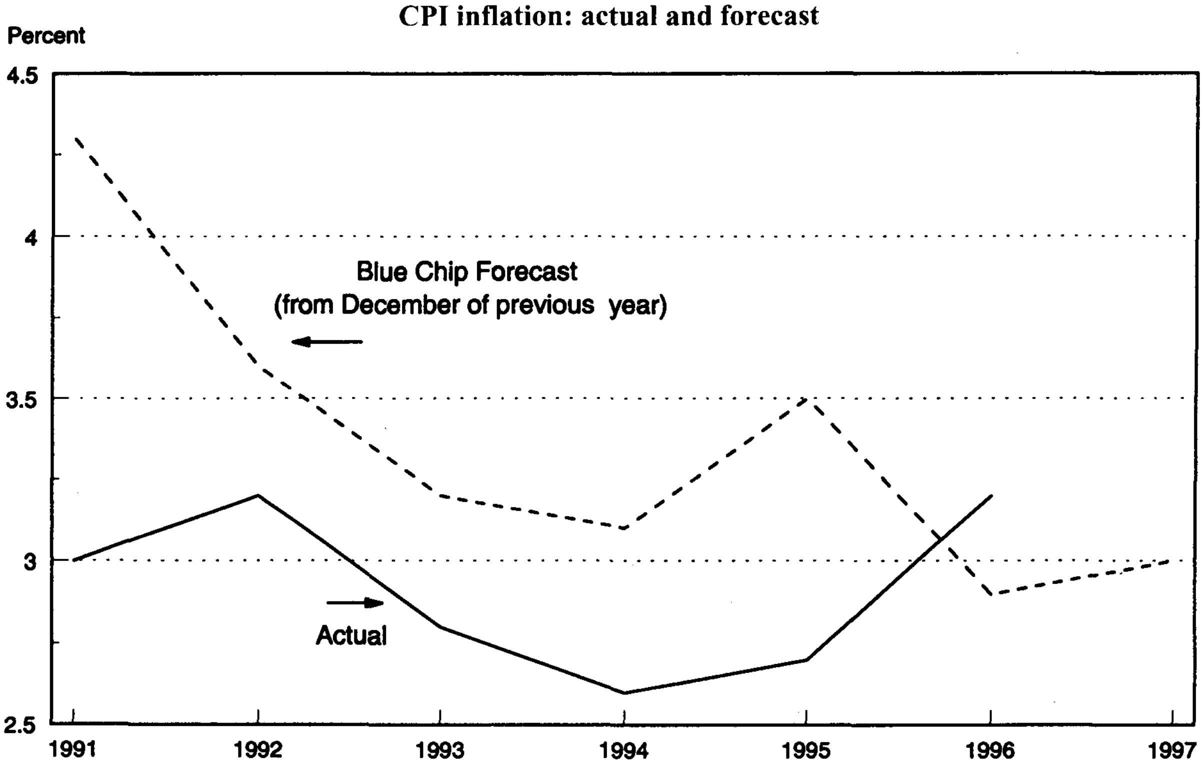


In a similar fashion, the lower panel of Chart 2 shows the rate of capacity utilization relative to a level of 83%. As Boldin (1995) notes, studies have generally associated accelerating inflation with capacity utilization rates in excess of 82-84%. While capacity utilization has since moved off its peak, it nevertheless remains quite high.

Consistent with these two indicators, the Blue Chip consensus forecast has, until recently, overpredicted inflation every year since the current expansion began (Chart 3). And, with the exception of this past year, the magnitude of the forecast errors continued to increase. Thus, using a wide variety of methods and models, forecasters also have been wrongly expecting an increase in the inflation rate.

While commentators have cited the type of evidence in Charts 1-3 to support their claims that inflation has remained unexpectedly low, these data only offer impressionistic evidence concerning a possible shift in the inflation process. To gain deeper insight into the recent behavior of inflation, one must examine the inflation process in a more formal manner. Following conventional practice, we interpret the movements in inflation as reflecting the influence of a set of key variables as well as various "shock" factors. While the underlying determinants are central to explaining the movements in inflation over longer periods, shocks to the inflation process can be viewed as exerting secondary effects which, at times, can fuel short-term bursts of inflation.

Chart 3
CPI inflation: actual and forecast

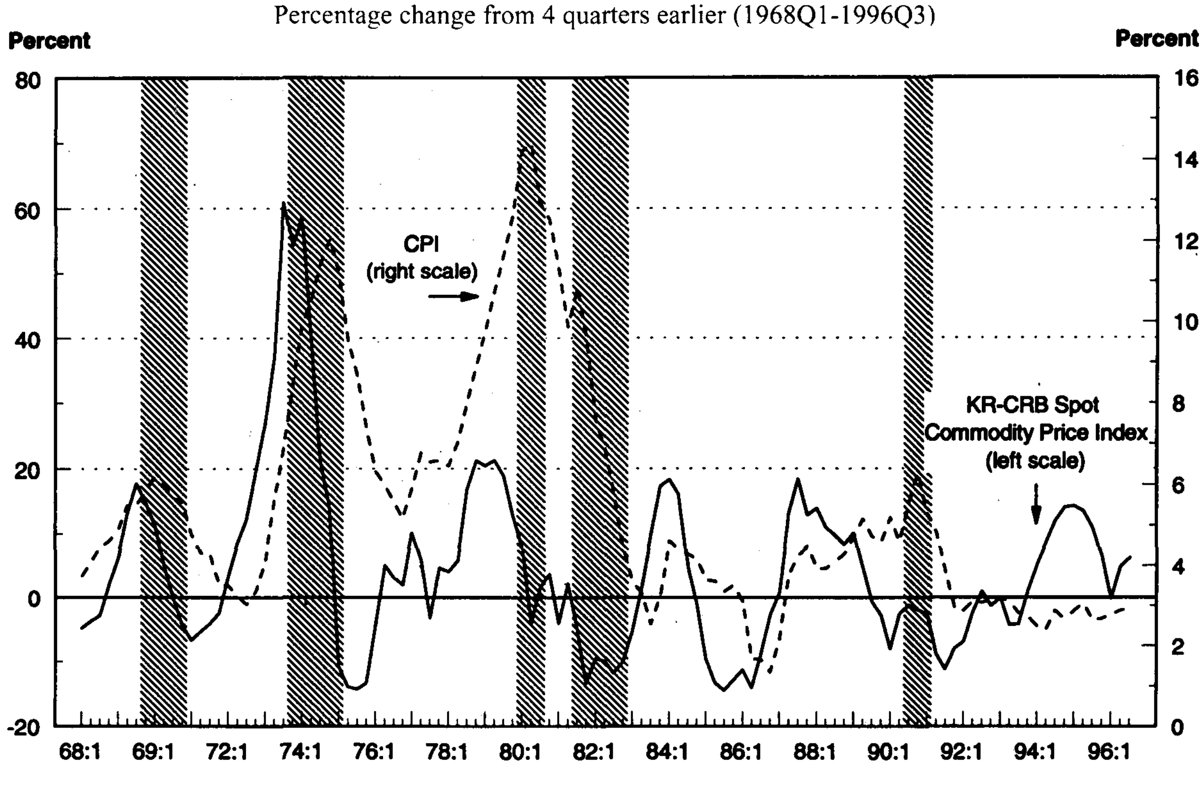


It is worth noting that some of these shock factors have managed to be well contained during the current expansion. One such factor is the absence of significant commodity price shocks. As Chart 4 shows, during most previous expansions, positive commodity price shocks occurred which added to inflationary pressures. However, recent evidence suggests that there is no longer a tight link between commodity prices and inflation. In particular, Blomberg and Harris (1995) document a marked decrease in the predictive power of commodity prices for inflation which they attribute to a decline in the commodity composition of US output beginning in the middle 1980s. Thus, while a commodity price increase occurred during this expansion, there are strong reasons to believe that its impact and contribution to inflation were considerably smaller than in previous expansions.

Chart 4

KR-CRB spot commodity price index and consumer price index

Percentage change from 4 quarters earlier (1968Q1-1996Q3)

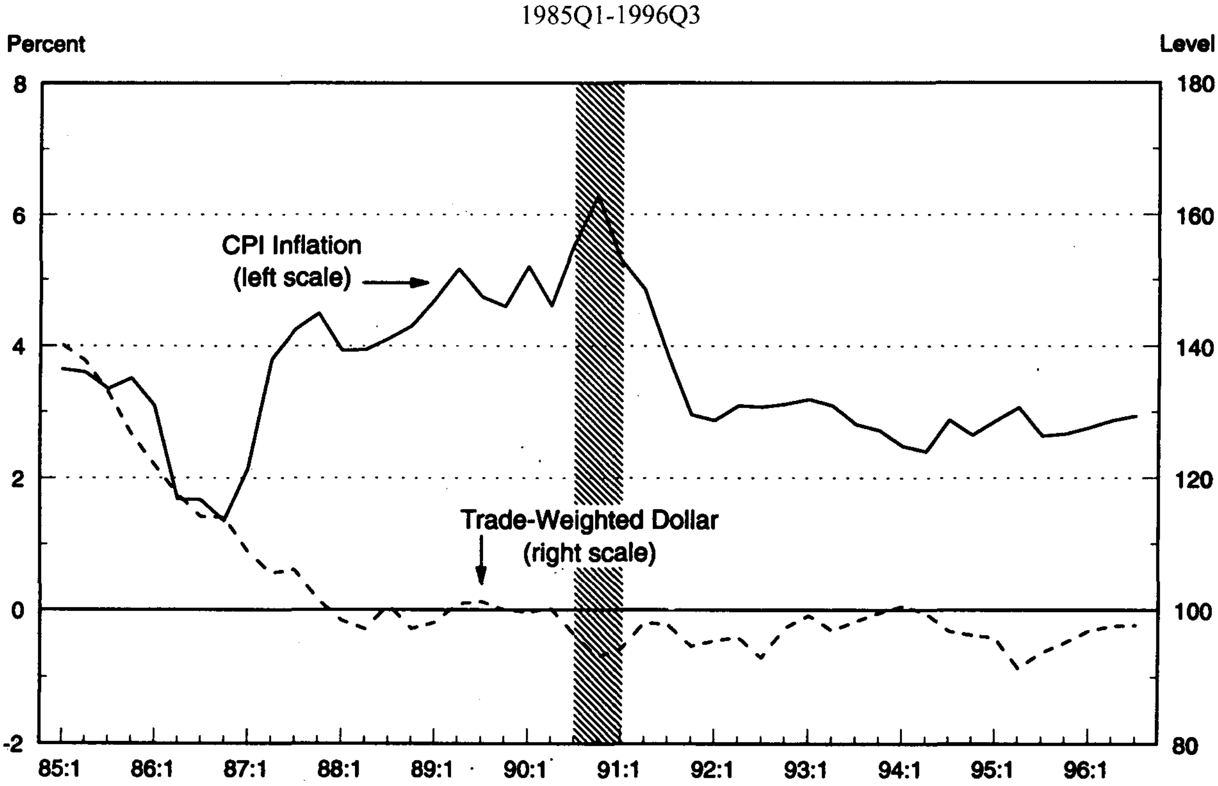


Note: Shading refers to NBER recessions.

Chart 5

CPI inflation and trade-weighted value of the dollar

1985Q1-1996Q3



Note: Shading refers to NBER recession.

A second factor concerns the behavior of the dollar. As shown in Chart 5, unlike the late 1980s when the 1985-87 dollar decline preceded the 1986-90 rise in inflation, the dollar has maintained its value in this expansion.

In the next section, we formulate a model to describe the movements in inflation over time. We briefly provide some background for our specification and report our estimation results. We then present several diagnostic tests, including out-of-sample forecasts, to determine how well our equation can account for the recent behavior of inflation.

2. A price-inflation Phillips-curve model

The "Phillips curve" refers to the posited relationship between the rate of change of a nominal wage or price and various indicators of real economic activity. The origin of the Phillips curve can be traced back to the work of Phillips (1958) who documented a strong inverse relationship between the rate of change of nominal wages and the level of unemployment in the United Kingdom. His findings were interpreted as establishing a wage adjustment process where low levels of unemployment represent tight labor markets that portend, or coincide with, an acceleration in wage growth. Subsequent versions of the Phillips curve recast the equation as a relationship between price inflation and unemployment, with the set of explanatory variables augmented to allow for the effects of expected inflation and other factors.[4] As Fuhrer (1995) notes, many of these subsequent additions were anticipated by Phillips in his original discussion.

In this section, we draw upon the Phillips-curve literature to specify and estimate a relationship between price inflation and its key determinants. We then use the estimated relationship to examine whether there is any evidence of a recent structural change in the inflation process. The price-inflation Phillips-curve model is given by:

$$INF_t = \alpha_0 + \alpha_1 GDPGAP_{t-1} + \alpha_2(\Delta GDPGAP_{t-1}) + \sum_{i=1}^{3}\alpha_{2+i}(INF_{t-i}) + \sum_{i=1}^{2}\alpha_{5+i}(OILG^*_{t-i}) + \sum_{i=1}^{2}\alpha_{7+i}(UNITG_{t-1}) + \varepsilon_t \quad (1)$$

where:

INF = inflation measured by the growth rate of the core CPI;

$GDPGAP$ = the output gap measured by the log ratio of actual to potential GDP;

$\Delta GDPGAP$ = the first difference (or change) in the output gap;

$OILG^*$ = net positive change in the real price of oil;

$UNITG$ = the growth rate of unit labor costs (nonfarm business sector);

ε = mean zero, serially uncorrelated random disturbance term.

Equation (1) uses the output gap (Chart 6) in place of the unemployment rate as a measure of aggregate demand, although the results are similar when the latter is used.[5] The idea of the effect of the output gap on inflation is similar to that of the unemployment gap: the economy operating above potential GDP is assumed to generate upward pressure on prices.[6] We also include

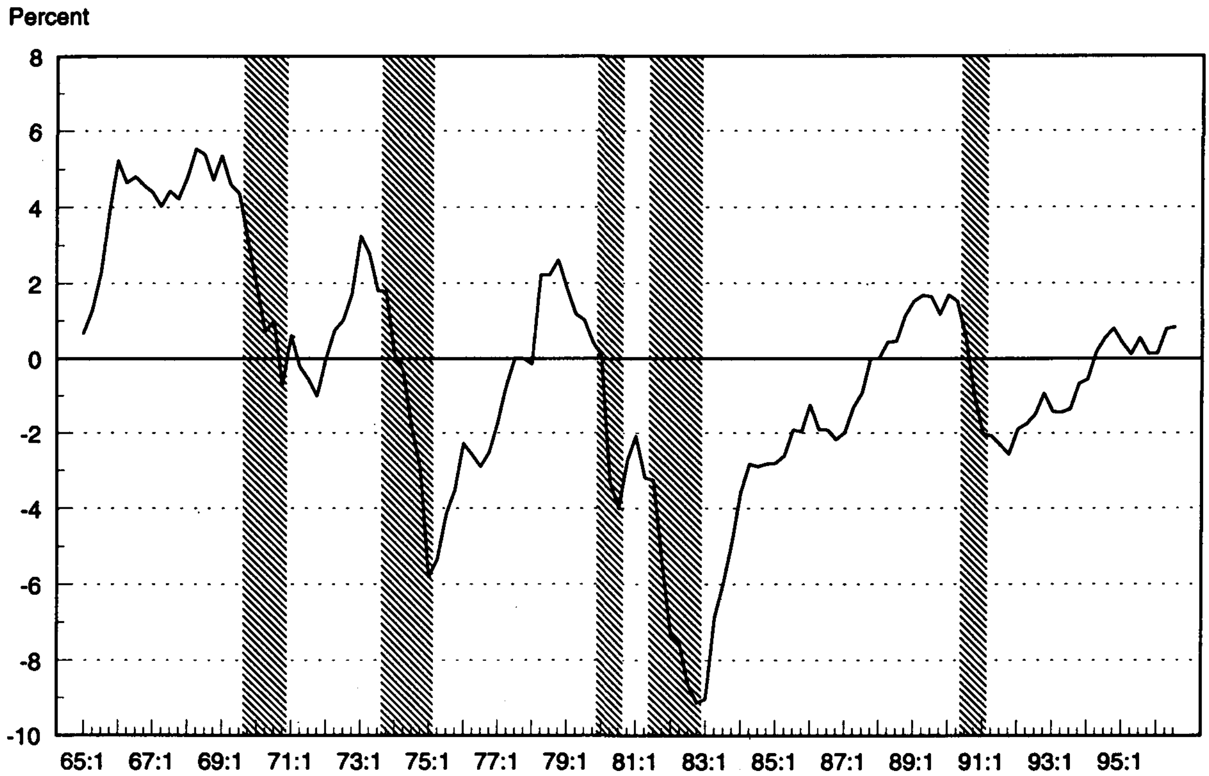
---

4 The theoretical development of the natural rate hypothesis and the distinction between the short-run and long-run Phillips-curve trade-off is due to the work of Friedman (1968) and Phelps (1967). There is a large empirical literature that has estimated current "expectations-augmented" Phillips curves. This literature includes the studies of Gordon (1970, 1975, 1977, 1982, 1990, 1996), Fuhrer (1995), Tootell (1994), King and Watson (1994), and King, Stock and Watson (1995).

5 More detailed definitions and sources of the data are presented in the Data appendix.

6 Fuhrer (1995) also uses the output gap as a measure of aggregate demand pressures.

Chart 6
The output gap
Percentage difference between actual and potential GDP (1965Q1-1996Q3)



Note: Shading refers to NBER recessions.

the first difference in the output gap to allow for a rate-of-change or "speed of adjustment" effect. More pressure is likely placed on prices when the gap narrows quickly rather than more slowly.[1]

The remaining basic determinants of inflation include past rates of inflation, oil prices, and unit labor costs. Lagged inflation terms are included to incorporate price inertia effects. Early researchers used past inflation rates to proxy expected inflation. In modern versions of the Phillips curve, Gordon (1996) has noted that such an interpretation is overly restrictive. In particular, he suggests that past inflation rates should be viewed as capturing the dynamics of price adjustment related not only to expectations formation, but also to the presence and extent of institutional factors in the economy such as wage and price contracts as well as delivery lags.

The model also allows for the influence of supply shocks. While changes in the relative price of inputs and the change in the real effective foreign exchange rate have been used as supply shock variables, we include a measure of the net positive change in real oil prices (Chart 7) in our specification. Because the core CPI excludes energy prices as a component, our supply shock variable attempts to capture any indirect channel of effect of oil prices on inflation. The construction of the supply shock variable follows from the approach of Hamilton (1996) and is designed to account for the change in the behavior of real oil prices and its increased volatility over the post-1986 period.[2]

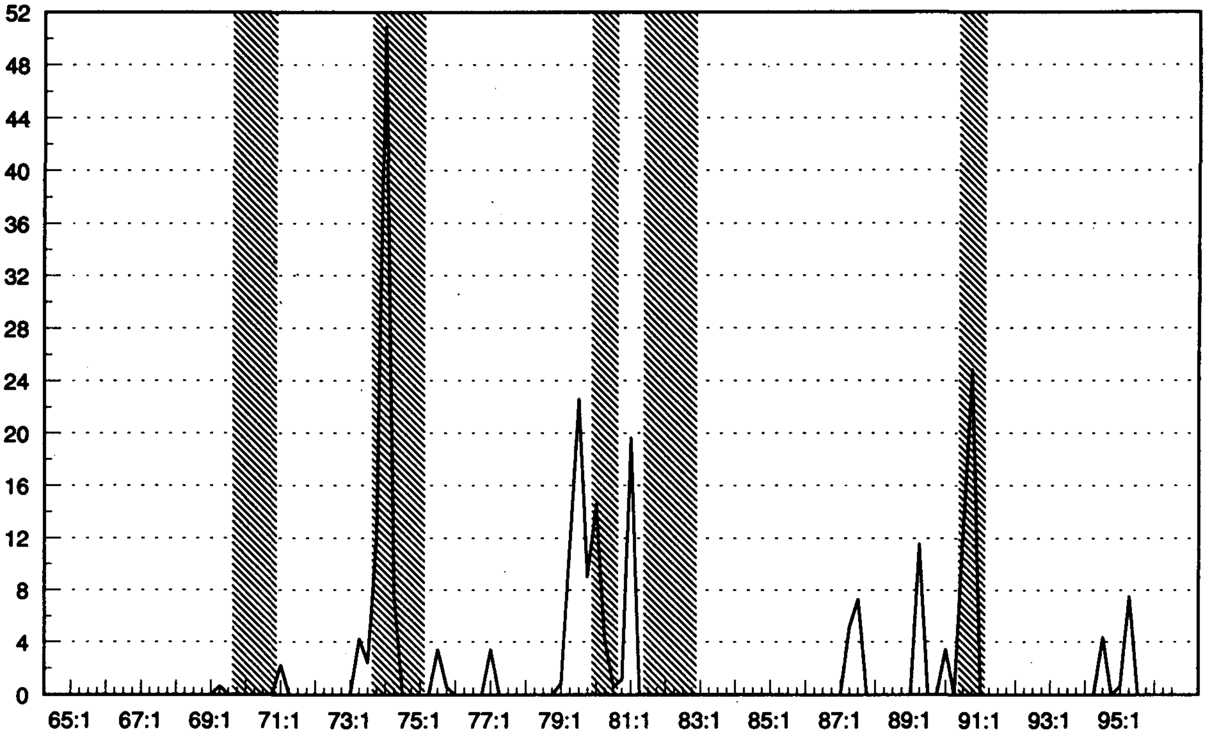
1 Gordon (1977, 1996) and Fuhrer (1995) have argued that rate of change effects for the output gap or unemployment gap are important for explaining the dynamics of the inflation process.

2 Net negative changes in real oil prices were excluded from equation (1), as they proved to be quantitatively and statistically insignificant.

Chart 7

Net positive change in real oil prices

1965Q1-1996Q3

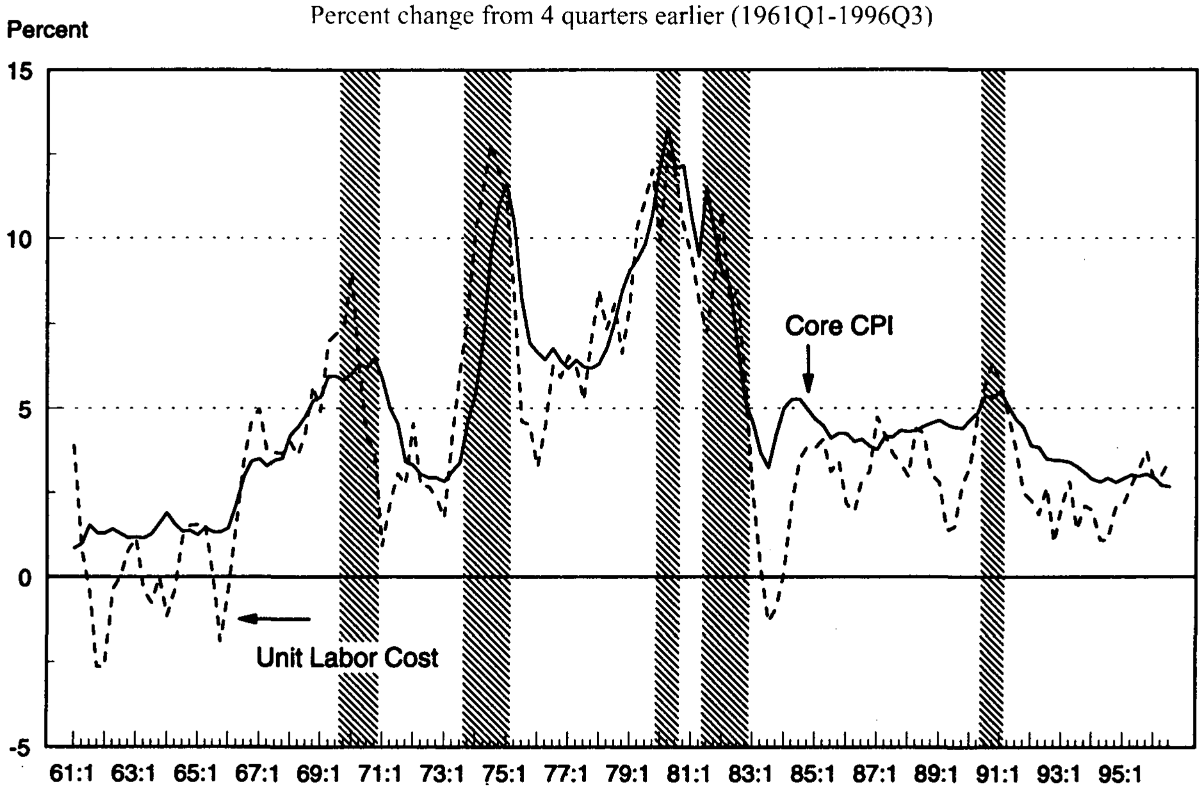


Note: Shading refers to NBER recessions.

Chart 8

Core CPI and unit labor costs

Percent change from 4 quarters earlier (1961Q1-1996Q3)

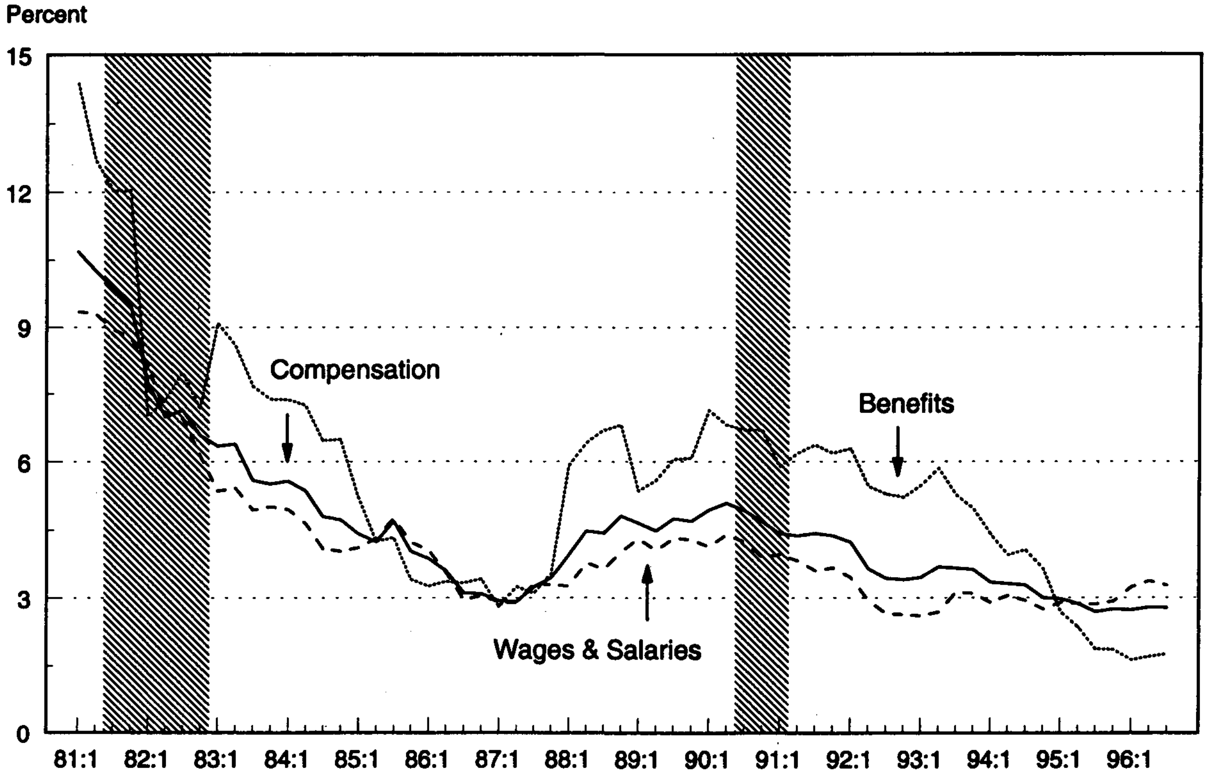


Note: Shading refers to NBER recessions.

The inclusion of the growth rate in unit labor costs (Chart 8), defined as compensation growth less productivity growth, allows for the notion that firms set prices as a markup over costs. Thus, when wage growth exceeds productivity growth, unit labor costs rise and so ultimately do prices. It is important to note that "traditional" price-inflation Phillips-curve models typically do not include unit labor costs as an explanatory variable. However, our specification allows us to account for the possibility that slow compensation growth during this expansion may have acted to offset other sources of inflationary pressures in the economy. Several commentators have suggested that the recent appearance of an inflation puzzle may be due to price-inflation Phillips curves neglecting to incorporate the effects of the decline in benefits growth and/or the restraint on wage growth that has occurred during this decade.[3] In particular, it is argued that these factors may have acted to lower the increase in overall labor costs and thereby reduced the pressure on firms to raise prices.

An examination of the behavior of compensation growth and its individual components (Chart 9) offers some support for this idea. Specifically, the growth rates for the three series are below the levels observed at the beginning of the current expansion, with benefits growth displaying the most dramatic decrease over this period.

Chart 9
Employment cost index: private industry
Percentage change from 4 quarters earlier (1981Q1-1996Q3)



Note: Shading refers to NBER recessions.

While the model takes real oil prices as exogenous, we only allow lagged values of the output gap and unit labor cost growth to be included as regressors to avoid simultaneity bias arising from the endogeneity of these variables. The lag lengths in equation (1) are selected by maximizing adjusted R^2, searching over one to four lags for inflation, the output gap, and unit labor cost growth

3 Meyer (1997) is among those who have recently discussed this point.

and zero to four lags for the net positive change in the real price of oil.[4]

Equation (1) is estimated by the method of ordinary least squares (OLS) using quarterly data over the period 1965Q1-1996Q3 and results are presented in Table 1. As the table shows, both the level of the output gap variable and its rate-of-change are highly significant and have the expected positive sign. The two lagged values of both unit labor cost growth and the net positive change in the real price of oil are also highly significant with the expected positive signs.

Table 1
1965Q1-1996Q3

$$INF_t = \alpha_0 + \alpha_1 GDPGAP_{t-1} + \alpha_2(\Delta GDPGAP_{t-1}) + \sum_{i=1}^{3}\alpha_{2+i}(INF_{t-i}) + \sum_{i=1}^{2}\alpha_{5+i}(OILG^*_{t-i}) + \sum_{i=1}^{2}\alpha_{7+i}(UNITG_{t-i}) + \varepsilon_t$$

Parameter	Estimate	Standard error	t-statistic	p-value
α_0	0.0430	0.0682	0.6303	0.5297
α_1	0.0196*	0.0098	2.0000	0.0478
α_2	0.2603**	0.0427	6.0961	0.0000
α_3	0.2681**	0.0778	3.4456	0.0008
α_4	0.1279	0.0841	1.5207	0.1310
α_5	0.2928**	0.0744	3.9334	0.0001
α_6	0.0169**	0.0047	3.5558	0.0005
α_7	0.0227**	0.0048	4.6934	0.0000
α_8	0.1842**	0.0415	4.4368	0.0000
α_9	0.0721*	0.0329	2.1886	0.0306

Adjusted R-squared	0.812	Chow forecast test	1992Q1-1996Q3
Standard error of regression	0.292	F-statistic	0.218 (0.999)
Number of observations	127	Log likelihood ratio	5.265 (0.999)
Q(30)	28.314 (0.554)		

Note: The Ljung-Box Q test statistic for serial correlation of the regression residuals is distributed asymptotically as chi-square with 30 degrees of freedom. Probability values for the test statistics are reported below in parentheses. * Significant at the 5% level. ** Significant at the 1% level.

It is also worth noting that the three lags of the inflation rate are generally significant, but that our estimated version of the price-inflation Phillips curve does not constrain the sum of the

---

[4] The compensation growth Phillips curve presented in Section 3 includes dummy variables to capture the effects from the imposition and relaxation of wage and price controls during the 1970s. These dummy variables are excluded from the price-inflation Phillips curve because they were statistically insignificant. Alternative dating schemes for the dummy variables (Gordon (1982)) also proved to be unimportant for explaining the dynamics of inflation during the 1971-75 period.

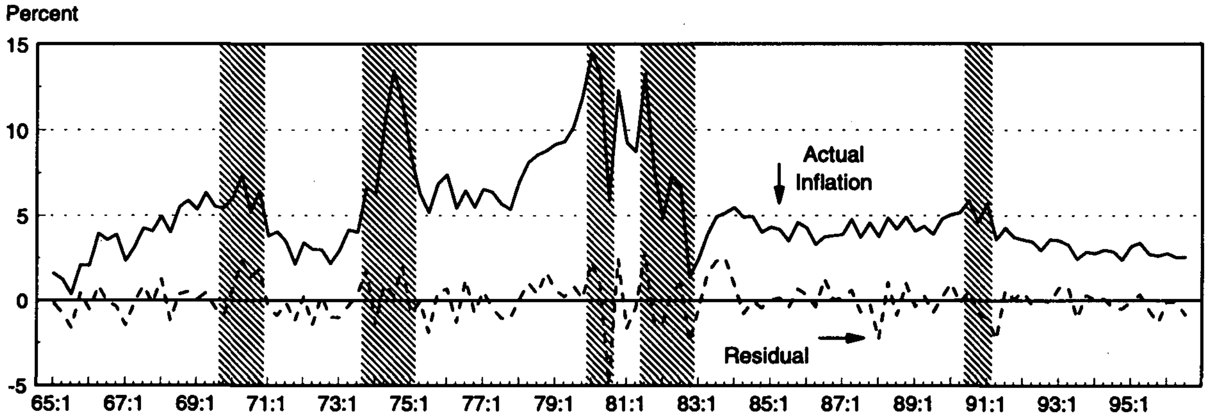
coefficients to equal unity ($\alpha_3+\alpha_4+\alpha_5=1$). This might initially suggest that equation (1) is inconsistent with a natural rate or accelerationist formulation of the Phillips curve. However, we will also present the estimation results of a compensation growth Phillips-curve model in Section 3. As shown in the Appendix, compensation growth can be solved out from the system of the two estimated equations to yield a reduced form of a price-inflation Phillips curve. The resulting model is characterized by coefficients on lagged inflation whose sum is not statistically different from unity and associates an acceleration in inflation with a positive output gap and a negative unemployment gap.

The adjusted R^2 indicates that the proportion of the variation of inflation that can be explained by the independent variables is quite high. We conducted several exercises to determine if any deterioration in the equation's fit had occurred in recent years. First, casual inspection of the estimated residuals plotted in the upper panel of Chart 10 shows no signs of deterioration.

Chart 10

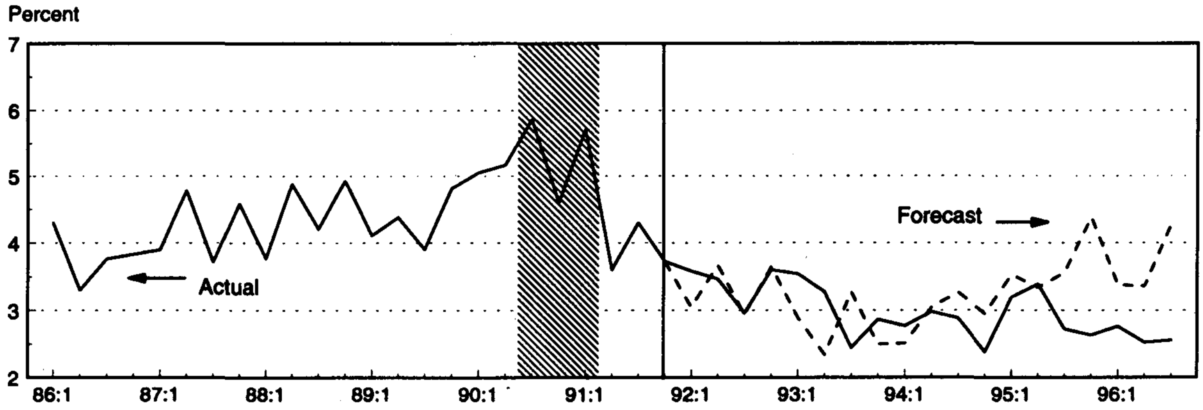
Core CPI inflation and within-sample residuals

1965Q1-1996Q3



Core CPI out-of-sample forecast

1992Q1-1996Q3



Note: Shading refers to NBER recessions.

We also applied Chow's (1960) split-sample test to the data to formally address the issue of parameter stability. There are several methods available to test the null hypothesis of constant parameters against the alternative hypothesis of a one-time shift in the parameters at some specified date. One method compares the estimates obtained using the data from one subsample (the beginning of the sample through the 1991 period) to the estimates using the full sample and yields a test statistic which is distributed asymptotically as F with $(m, n\text{-}k)$ degrees of freedom under the null hypothesis.[5] Another method uses dummy variables for the entire parameter vector for one subsample (the post-1991 period) and then tests the joint significance of the dummy variables. This latter method yields a test statistic that is distributed asymptotically as chi-square (χ^2) with k degrees of freedom under the null hypothesis. As shown by the values of these two test statistics reported in the table, we fail to reject the null hypothesis of parameter stability for the post-1991 period at conventional significance levels.[6]

We examined the equation's out-of-sample forecast performance as well. This dynamic simulation differs from the previous (within-sample) estimation exercise by using lagged forecasted values of inflation, rather than the lagged actual values of inflation, to generate the subsequent quarter's forecast. For this part of the analysis, we estimated the equation through 1991Q4 and then used the estimated equation to forecast inflation over the 1992-96 period. The lower panel of Chart 10 shows this forecast along with the actual values of inflation. As the chart indicates, the equation was fairly accurate in forecasting inflation through the middle of 1995. Over the last year, however, some large discrepancies between actual and forecasted inflation have occurred. In particular, it appears that an oil price shock in 1995Q2 was the principal contributor to the sizable overprediction of inflation in 1995Q3. Because the model is being used to generate dynamic forecasts, the prediction error in 1995Q3 continues to affect the subsequent quarters' forecasted values. We performed the same exercise over the 1994-96 period with similar results (not shown).[7]

Having established that the relationship between inflation and its historic proximate determinants has remained relatively stable over this expansion, we next examine the key variables in the inflation equation and their relevance as possible sources of the slow growth in prices.[8] The two obvious contenders are the output gap variable and unit labor cost growth. We explore each of these two variables in turn.

The relatively low inflation rates experienced over the current expansion could have been the result of slow and steady output growth leading to only a gradual narrowing of the gap between actual and potential output (Chart 6). Whereas in previous expansions this gap variable rose to 2% or more, in this expansion it has barely remained above zero. This fact suggests that the lack of "hot" quarters of economic growth might be playing a role in keeping inflation low.

---

5 The values of n and $n+m$ refer to the number of observations in the first subsample and the total sample, respectively. The value of k refers to the number of parameters in the model.

6 We also examined the data for evidence of parameter instability using the CUSUM and CUSUMSQ tests proposed by Brown, Durbin and Evans (1975). The tests are based on recursive residuals, with the CUSUM test primarily used for detecting gradual structural change and the CUSUMSQ test for sudden structural change. The tests provided no evidence of parameter instability and corroborate the previous results that the low rates of inflation observed during this expansion are not indicative of any structural change.

7 One possible explanation for the recent divergence between actual and forecasted inflation is the discrepancy, since mid-1994, between the income and product sides of the National Accounts. If the product side is revised up to align better with the income side, as some have suggested, then productivity will also be revised up and both the output gap and unit labor costs will be revised down. These latter revisions would lower our out-of-sample forecast for the 1995-96 period. (See Macroeconomic Advisers, LLC, 1996 for a discussion of the discrepancy in the National Accounts.)

8 Note that our model does not allow us to examine whether there has been a shift in the Federal Reserve's inflation fighting credibility that could have changed the inflation process by directly altering inflationary expectations. Such an examination is beyond the scope of this paper: it would involve estimating a separate equation for inflationary expectations with some measure of Fed credibility included as an explanatory variable. However, the stability of the equation's structure strongly suggests that such a shift has not taken place. Blanchard (1984) has noted that Phillips curves of this type remained stable even after the 1979 change in Federal Reserve operating procedures.

The problem with the preceding argument is, however, that while it might explain why inflation has not accelerated, it does not account for the slow *decline* in the inflation rate that has occurred in more recent years. As can be seen in Chart 6, the difference between actual and potential output was not all that large at the start of the current expansion. In addition, it has not taken any longer for the gap variable to turn positive in this expansion relative to others. Thus, we do not believe that relatively modest output growth has played the principal role in the inflation story.

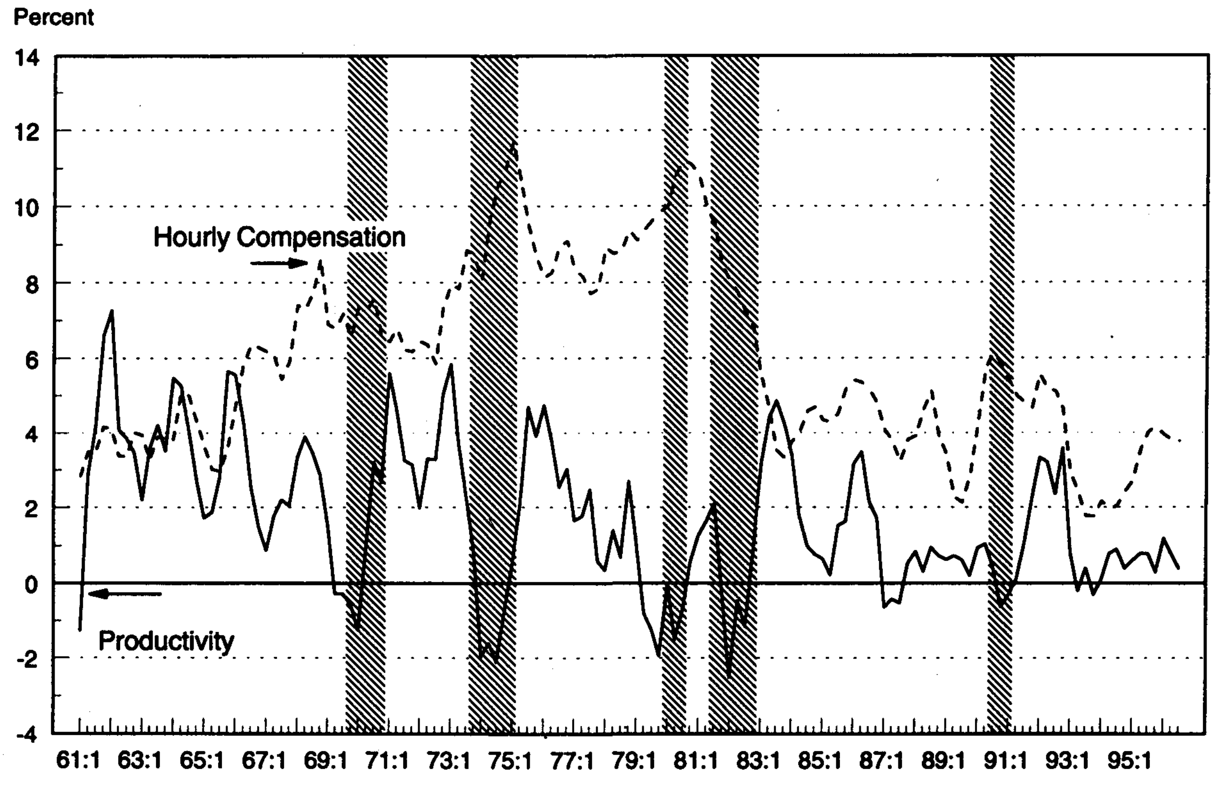
The second variable to consider is unit labor cost growth. As Chart 8 shows, not only has growth in unit labor costs been weak during this expansion, but a gap between unit labor cost growth and CPI inflation opened up during most of the expansion. Only recently has this gap narrowed. Thus, an initial examination of the data appears to be consistent with the view that unit labor costs may have played an important role in restraining the pressure on firms to raise prices.

With unit labor costs defined as compensation (wages and benefits) divided by productivity, either slow compensation growth or fast productivity growth must be accounting for its weak growth. As Chart 11 shows, productivity growth has not been unusually strong in the current expansion. During late 1991 and early 1992, the series rose at roughly a 3% rate contributing to weaker growth in unit labor costs. But since that time, productivity has grown at rates below 1%. The chart also shows that compensation growth declined to around 2% fairly early in the expansion and hovered around that rate for over a year before showing signs of a modest pick-up. This 2% growth rate is below any rate recorded over the past 35 years, suggesting that compensation growth appears to be responsible for the slow growth in unit labor costs. In the next section, we examine compensation growth and its recent behavior in more detail.

Chart 11

Productivity and hourly compensation

Percentage change from 4 quarters earlier (1961Q1-1996Q3)



Note: Shading refers to NBER recessions.

3. Compensation growth Phillips curve

In an exercise that parallels the analysis in Section 2, we present the results from estimating a wage-inflation Phillips-curve model for compensation growth. We then examine whether there is evidence to suggest a recent change in the fit of the model to the data.

As previously noted, the original Phillips curve focused on the relationship between the change in nominal wages and the level of unemployment. For present purposes, we expand the dependent variable, wages, to also include benefits. In recent years benefits have become an increasingly important part of workers' compensation. The compensation growth Phillips-curve model is given by:

$$LXNG_t = \beta_0 + \sum_{i=1}^{2}\beta_i\left(LXNG_{t-i}\right) + \beta_3 U_{t-1} + \sum_{i=1}^{2}\beta_{3+i}\left(INF_{t-i}\right) + \beta_6 SOC_t + \beta_7 UIR_{t-1} + \beta_8 DUM_t + \eta_t \quad (2)$$

where:

$LXNG$ = growth rate of compensation per hour (nonfarm business sector);

U = unemployment rate for males 25-54 years old;

INF = inflation measured by the growth rate of the CPI (all items, urban consumers);

SOC = the change in employer social security contributions;

UIR = income replacement ratio from unemployment insurance benefits;

DUM = dummy variable for the Nixon wage and price controls;

η = mean zero, serially uncorrelated random disturbance term.

Equation (2) principally relates the movements in compensation growth to the unemployment rate and other variables reflecting labor market conditions.[9] The unemployment rate of prime age males is used as a measure of labor market tightness. We enter the variable in levels and thereby abstract from any explicit discussion of the NAIRU other than to note that the specification can be viewed as implicitly assuming a constant value for the NAIRU over the sample period.[10] Equation (2) does not include a rate-of-change effect for the unemployment rate because the estimated coefficient on a second lag of the unemployment rate was quantitatively and statistically insignificant and, therefore, was omitted from the specification.[11]

The remaining determinants of compensation growth include the change in employer social security tax contributions which is a component of hourly compensation. The income replacement ratio from unemployment insurance benefits attempts to capture changes in compensation growth related to job search. A dummy variable accounts for the restraining effect of the price freeze in 1971Q4 and the rebound effect after the relaxation of the controls in 1972Q1.[12] The inclusion of lagged inflation terms parallels the previous discussion concerning wage and price inertia. Last, we only allow lagged values of the unemployment rate and inflation rate to be included as regressors because of endogeneity considerations.

---

9 Detailed definitions of the data and their sources are reported in the Data appendix.

10 For example, we could follow the approach of Fuhrer (1995) who assumes a value of 6% for the NAIRU and use the unemployment gap (the difference between the actual level of unemployment and the NAIRU) instead of the unemployment rate as an explanatory variable in equation (2). However, this will not affect the regression results other than to change the estimated values of the constant term and the coefficient on the unemployment rate.

11 This result is consistent with Fuhrer (1995) who also finds no significant rate-of-change effects for the unemployment rate in wage-inflation Phillips-curve models.

12 The definition of the dummy variable is from Englander and Los (1983).

Equation (2) is estimated using quarterly data over the period 1967Q2-1996Q3 and the OLS results are presented in Table 2. As the table indicates, the first lagged values of the dependent variable and price inflation are not significant at the 5% level, while these variables' second lags are both significant. The unemployment rate is significant and has the expected negative sign. Finally, the variables reflecting labor market conditions are all significant with the expected signs. The adjusted R^2, although not quite as high as the value reported in Table 1, also indicates that the estimated equation fits the data quite well. The Chow break test does not reject the null hypothesis that the discrepancies between the actual and predicted values are statistically insignificant. However, inspection of the equation's residuals depicted in the upper panel of Chart 12 reveals some evidence of a modest break beginning in the early 1990s.

Table 2

1967Q2-1996Q3

$$LXNG_t = \beta_0 + \sum_{i=1}^{2} \beta_i (LXNG_{t-i}) + \beta_3 U_{t-1} + \sum_{i=1}^{2} \beta_{3+i}(INF_{t-i}) + \beta_6 SOC_t + \beta_7 UIR_{t-1} + \beta_8 DUM_t + \eta_t$$

Parameter	Estimate	Standard error	t-statistic	p-value
β_0	0.4247	0.2299	1.8476	0.0674
β_1	0.1591	0.0850	1.8713	0.0640
β_2	0.1854*	0.0830	2.2326	0.0276
β_3	-0.0662**	0.0243	2.7188	0.0076
β_4	0.1599	0.0879	1.8201	0.0715
β_5	0.1863*	0.0885	2.1044	0.0376
β_6	0.0905**	0.0206	4.4232	0.0000
β_7	1.6440*	0.8223	1.9994	0.0481
β_8	-0.7396*	0.3419	2.1633	0.0327

Adjusted R-squared	0.681	Chow forecast test	1992Q1-1996Q3
Standard error of regression	0.391	F-statistic	0.722 (0.787)
Number of observations	118	Log likelihood ratio	16.749 (0.607)
Q(30)	20.778 (0.895)		

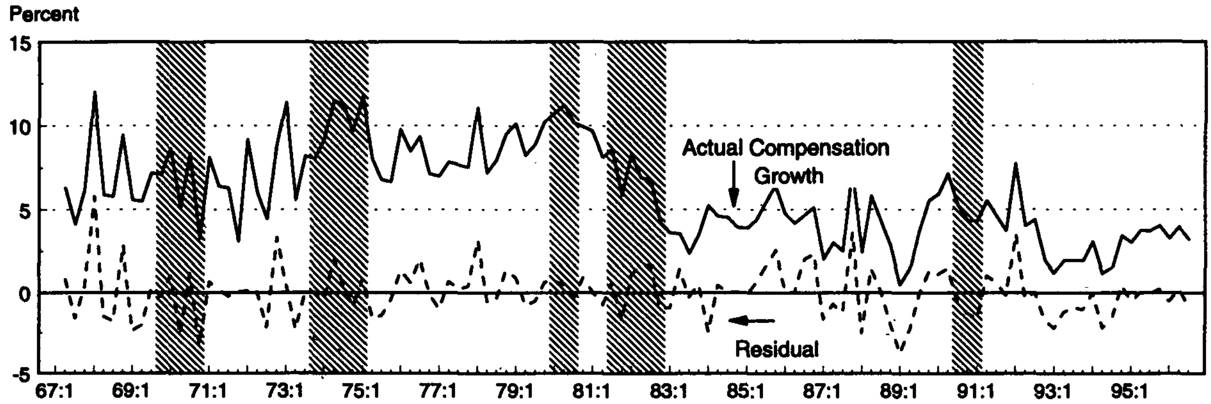
Note: The Ljung-Box Q test statistic for serial correlation of the regression residuals is distributed asymptotically as chi-square with 30 degrees of freedom. Probability values for the test statistics are reported below in parentheses. * Significant at the 5% level. ** Significant at the 1% level.

To gain further insight into the issue of a possible structural break, we estimate equation (2) over the period 1967Q2-1991Q4 and then use this estimated equation to generate predicted values for compensation growth over the 1992-96 period. When the estimated equation is used to generate a dynamic out-of-sample forecast, the model appears to consistently overpredict compensation growth from roughly 1992 until 1994 (lower panel of Chart 12). Based on the past relationship with its explanatory variables, compensation growth should have been higher by roughly 2% during this time period. Further, the timing and magnitude of this shortfall appears to be consistent with the observed breakdown of various indicators (Charts 1-3) that signaled an imminent acceleration in inflation and provides additional support for the view that unit labor costs may be the key element in understanding the origin of the inflation puzzle. Although compensation growth now appears to be "back on track", and price-inflation was never "off-track", it is still of interest to try to understand the factor(s) responsible for the temporary shortfall in compensation growth.

Chart 12

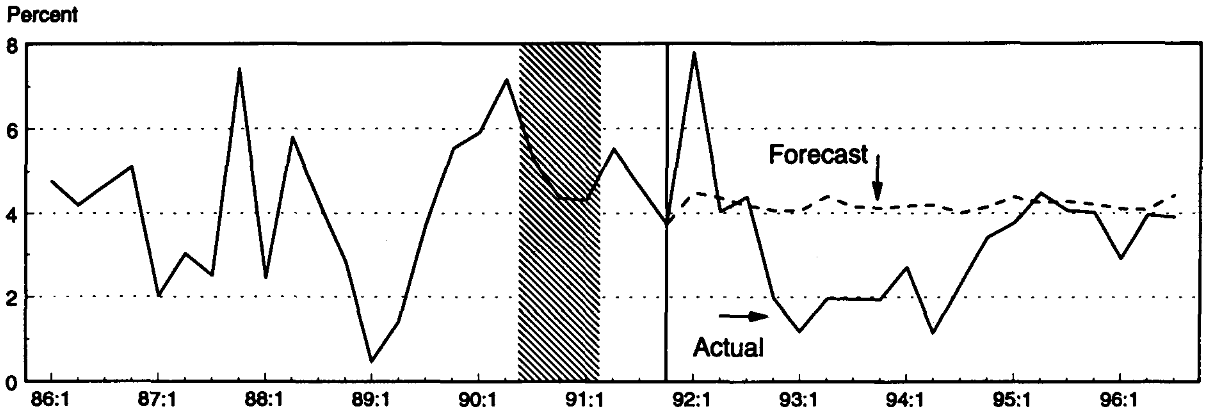
Actual compensation growth and within-sample residuals

1967Q1-1996Q3



Compensation growth out-of-sample forecast

1992Q1-1996Q3



Note: Shading refers to NBER recessions.

Countless numbers of newspaper articles, along with several more rigorous pieces, have been written discussing whether changes in the behavior of labor markets have served to mitigate inflationary pressures during the current expansion.[13] One view is that labor markets are not as tight as traditional measures such as the unemployment rate might indicate. For our purposes, this explanation suggests that if an alternative labor market variable can be identified and included in our regression equation, then the 1992-94 overprediction of compensation growth can be greatly reduced or possibly eliminated. A second view is that a real or imagined decline in workers' power has occurred which is serving to reduce workers' willingness to ask for higher wages. Among the factors cited as support for this theory is the decline in the number of union members, the increase in the number of contingent workers, or in the number of firms moving their operations abroad where labor

[13] See for example Bradsher (1995), Passell (1995), Spiers (1995, 1996) and Valletta (1996).

is relatively cheap. Once again, our ability to assess the validity of this explanation will depend on the extent that we can quantify these particular phenomena. In the next section, we examine a number of series that attempt to measure these possible changes in labor market behavior.

4. Possible factors underlying compensation growth shortfall

To provide a further investigation into the 1992-94 compensation growth shortfall and the underlying causes, we proceed in two steps. First, we conduct an informal analysis. Specifically, we consider whether the movements in a particular series appear different in this expansion and could possibly explain the recent shortfall in compensation growth. If the informal evidence suggests that a particular series may be playing a role, then we turn to a more formal analysis. In particular, the second step involves testing whether the series adds significant explanatory content to our compensation growth equation as well as examining its quantitative importance for the recent weakness in compensation growth.

Admittedly, this second step is a difficult hurdle. Some of the factors may not have mattered in the early years of our sample period, and therefore might not display statistical significance over the entire period. On the other hand, if we only consider the effects of a variable during the 1990s, then this approach could potentially generate a significant correlation that is spurious in nature. Consequently, our empirical analysis will only consider the predictive content of variables based on the entire sample period. However, we recognize that any conclusions must be interpreted as suggestive, not definitive, and that the informal evidence might need to be given more weight than is customary.

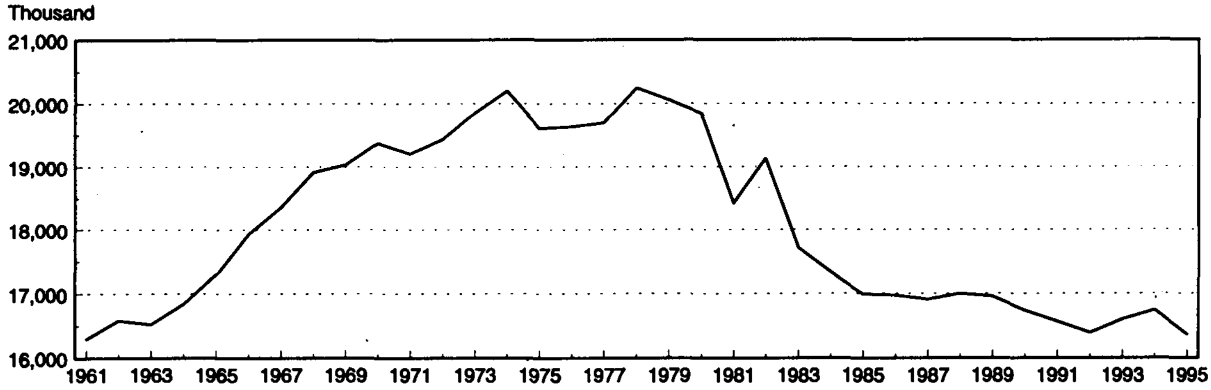
The labor market measures we examine are: union membership, the number of contingent or temporary workers, the average duration of unemployment, the number of job losers, consumer confidence concerning job prospects, and the number of job leavers. Each of these measures has been either explicitly or implicitly linked to weak compensation growth in one or more articles. We now discuss each of these measures in turn.

Decline in union membership. The upper panel of Chart 13 presents the number of union members since the early 1960s. As the chart indicates, union membership peaked in the early 1970s and has been declining ever since. The fact that this series has been declining for 20 years and has recently leveled off makes it hard to believe that a recent weakening of union power is responsible for the somewhat unexpected softness in compensation growth during the 1990s. The only basis for believing that the decline in union membership is playing a role is if its importance is only now evident because of a threshold effect, i.e. a decline in union membership only matters when it moves below some critical level. But such a hypothesis is difficult to test empirically. Related measures, such as the number of work stoppages occurring each year, have behaved in a similar fashion. Thus, the behavior of this series makes it doubtful that it can explain the 1992-94 overprediction in compensation growth.

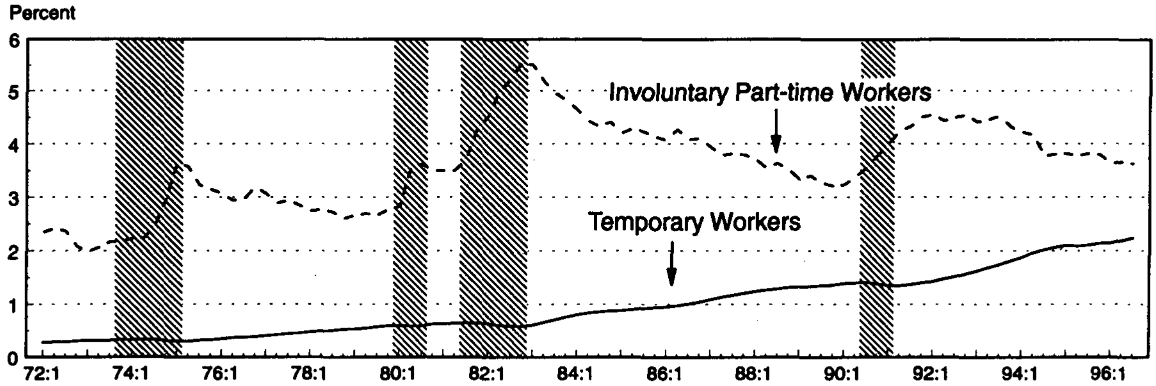
Contingent workers. A second possibility is that the growth in the number of contingent workers, those hired on a temporary basis who presumably have little or no bargaining power, has served to keep compensation growth weak. There are three categories of contingent workers: temporary workers, involuntary part-time workers and the self-employed. Because of the difficulty in separating the temporary from the permanent self-employed, our discussion focuses on only the first two categories.

The lower panel of Chart 13 presents temporary workers and involuntary part-time workers as a percent of nonfarm payroll employment. As the chart shows, the percentage of the labor force that consists of temporary workers has been growing steadily since the early 1970s. Yet even today, temporary workers make up only 2% of the workforce. This low percent makes it hard to believe that growth in temporary workers can account for the recent weakness in compensation

Chart 13
Union members
1961-1995



Temporary and involuntary part-time workers: percentage of nonfarm employment
1972Q1-1996Q3



Note: Shading refers to NBER recessions.

growth. Moreover, given the behavior of the series, one would again have to rely on a threshold effect for this series to account for the overprediction of compensation growth.

The involuntary part-time workers series behaves countercyclically, rising during recessions and falling during expansions. This category of workers was larger during the 1981-82 recession than it was during the 1990-91 recession, suggesting that these workers are not the source of any recent change in labor market behavior. Relative to its previous behavior, however, this series has been falling at a slower rate in this expansion. This slower decline could be playing some role in keeping wage growth modest.

When the involuntary part-time workers series is included in our compensation growth equation, the estimated coefficient has the expected negative sign, but it is not statistically significant. Thus, we are not able to find solid evidence that contingent workers are playing a key role in the shortfall in compensation growth during the early 1990s.

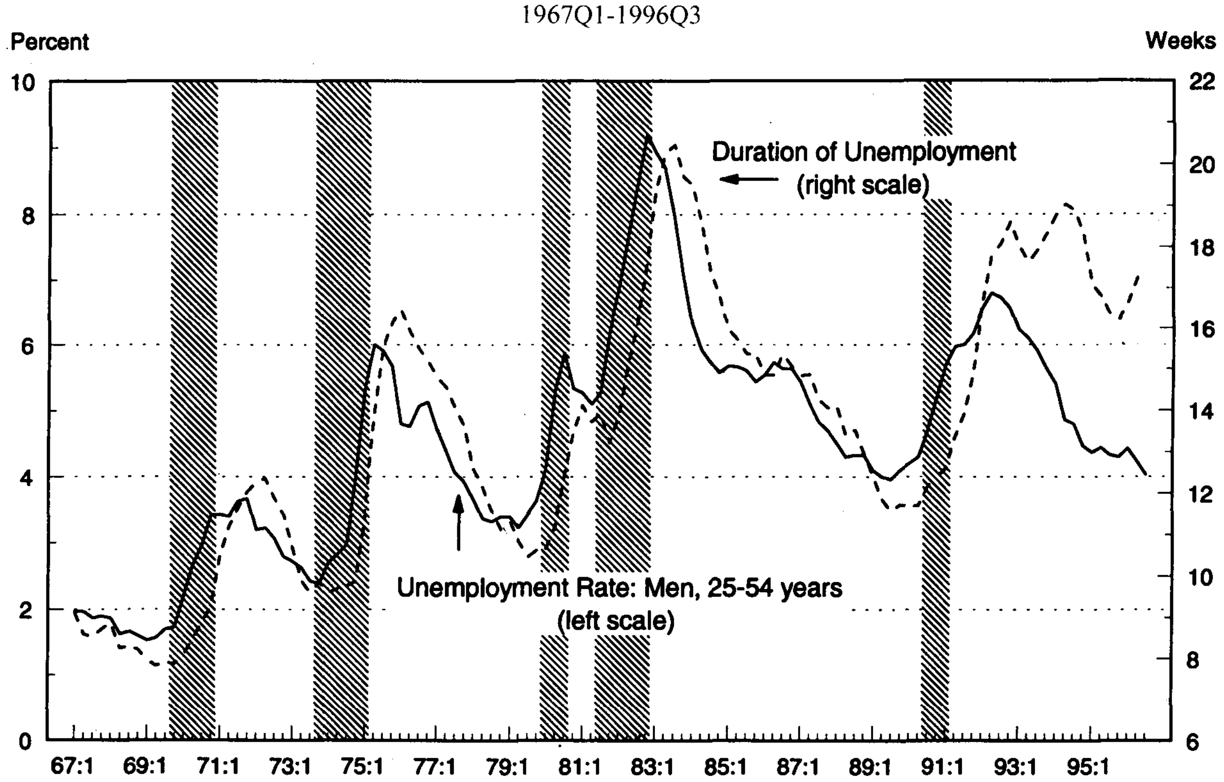
Average duration of unemployment. Some articles have claimed that the average duration of unemployment provides a better measure of slackness in labor markets than is suggested by the unemployment rate.[1] Some articles also have argued that this increase in unemployment duration has contributed to worker anxiety. As shown in Chart 14, the average duration of unemployment has

---

[1] See Sullivan (1996) for a discussion of this issue and some evidence that suggests the data do not support such a claim.

behaved somewhat differently in this expansion. In previous business cycles, the peak of this variable typically coincided with the start of an expansion. In this expansion, average duration remained near its peak for several years and only recently has begun to fall. Further, as the chart shows, the duration of unemployment has been high in this expansion relative to the unemployment rate. Thus, at least impressionistically, the average duration of unemployment appears to reflect some change in labor market behavior.

Chart 14
Average duration of unemployment and the unemployment rate
1967Q1-1996Q3



Note: Shading refers to NBER recessions.

Adding unemployment duration as a regressor to our compensation growth equation is problematic because of its high correlation with the unemployment rate (correlation 0.82). This variable is significant when the unemployment rate is excluded, although no more significant than the unemployment rate. Most importantly, the inclusion of unemployment duration in the equation does nothing to eliminate the 1992-94 compensation growth shortfall.

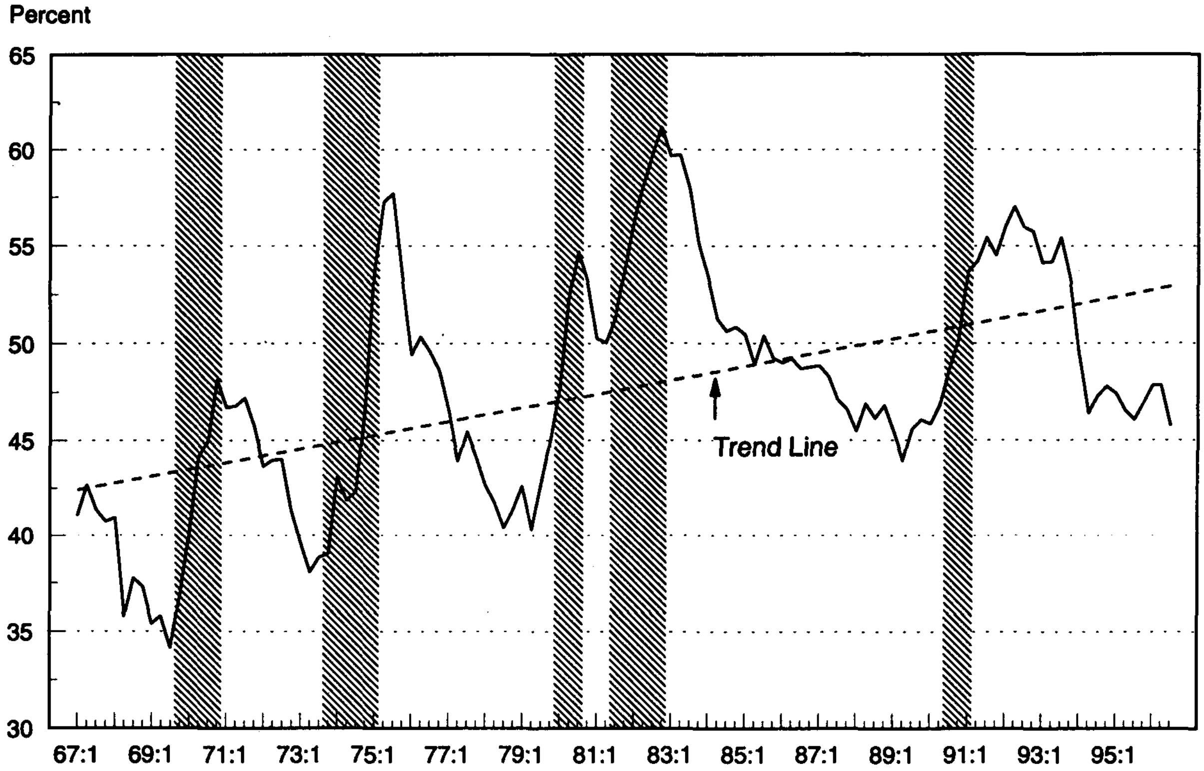
Job losers. Another variable that might be affecting workers is job security. If job security has fallen, workers might be less willing to ask for increases in compensation. Valletta (1996) argues that to examine this issue one should examine job dismissals (job losers plus temporary job completers as a percent of total unemployment). He notes, as Chart 15 indicates, that there has been a positive trend in permanent job dismissals, indicating that job security has decreased. But with the most recent data included, this trend seems to be less apparent in the 1990s. However, the series did not decline at the start of the expansion, as it had in previous expansions, so perhaps job security has been relatively worse in this expansion.

When job dismissals are included as an additional variable in our compensation growth equation, we find it is not statistically significant. The job dismissal series is significant when the unemployment rate is excluded, although no more significant than is the unemployment rate. Further, replacing the unemployment rate with the job dismissal series does nothing to improve the equation's out-of-sample forecasting ability. Overall then, the job dismissal data provides no additional information for compensation growth.

Chart 15

Job losers and temporary job completers as a percent of total unemployed

1967Q1-1996Q3



Note: Shading refers to NBER recessions.

Consumer confidence concerning job prospects. In addition to increases in the average duration of unemployment and declines in job security measures, some observers have argued that continued corporate restructuring and talk of companies moving abroad have raised workers' fears of job loss. This increased fear could be serving to keep compensation growth in check. To consider such a linkage, we examine an index of consumers' views about future job prospects – what might be called a consumer apprehension index.[2] This is shown, along with the unemployment rate, in Chart 16.

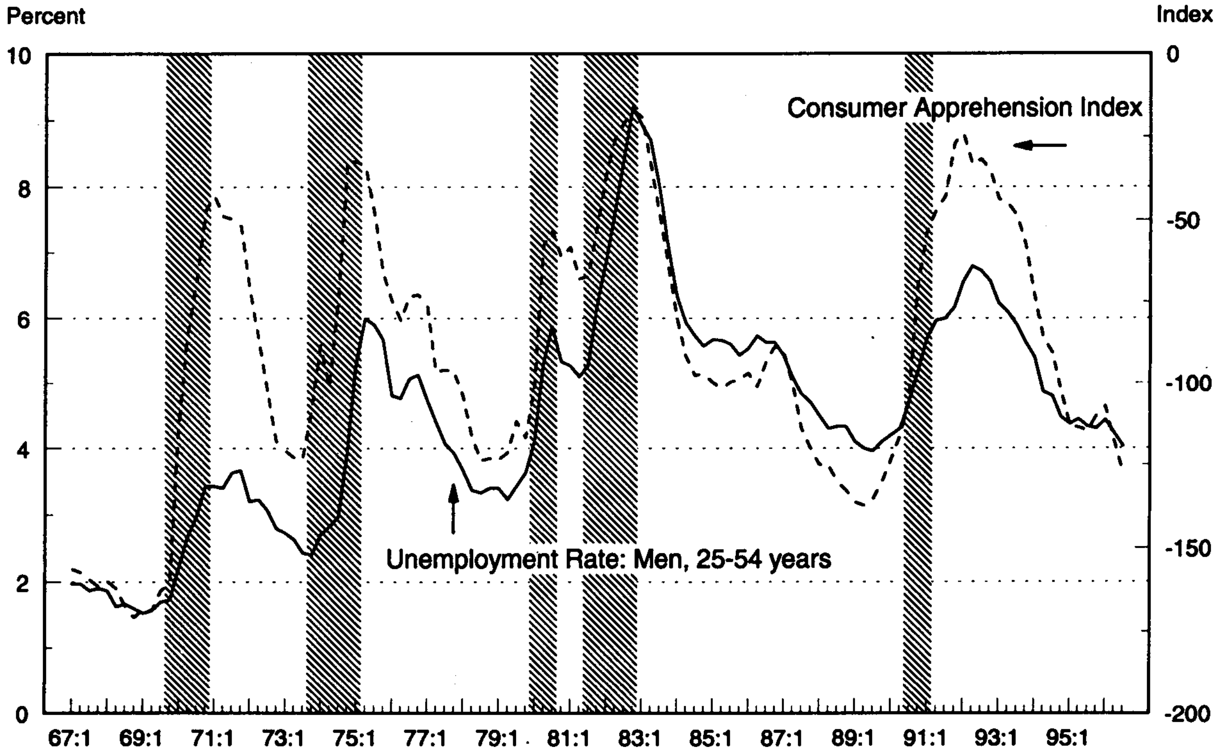
As the chart shows, the recent behavior of this index differs from its behavior in previous expansions: the index has been quite high relative to the observed unemployment rate. The only other time the relationship between the two series looked similar was during the years 1971-73, and at that time the gap between them narrowed much faster. However, much like the other variables, this index provides no additional information about compensation growth when it is included in the regression equation, nor does it improve the out-of-sample forecast.

Job leavers. The one measure that seems to help modestly in explaining compensation growth is job leavers as a percent of the civilian labor force. This variable measures the number of workers who have chosen to leave their jobs. If workers now have increased concerns about the

---

2 This series is constructed as the negative of the present employment situation component of the Consumer Confidence Index. For further detail see the Data appendix.

Chart 16
Consumer apprehension concerning job loss and the unemployment rate
1967Q1-1996Q3



Note: Shading refers to NBER recessions.

robustness of the labor market, one might expect that workers would be less likely to quit their present job. As the upper panel of Chart 17 indicates, this series has also behaved somewhat differently in this expansion and is currently at levels not seen since the 1960s. When included in our regression, this variable proves to be statistically significant at the 10% level. However, as shown in the lower panel of Chart 17, the series brings the out-of-sample forecasts only slightly closer to the series' actual values.

In summary, while numerous accounts have suggested that recent changes and developments in the labor market have affected compensation growth, we are able to find little confirmation of such a change. The impressionistic evidence suggests that there are differences in the behavior of some labor market variables in this expansion, but these differences appear to be statistically and economically unimportant. With the exception of job leavers, no labor market variable plays a significant role in explaining compensation growth. Moreover, the job leavers series does not lead to an appreciable improvement in the out-of-sample forecast of compensation growth during the early 1990s. While it is possible that labor market changes are occurring that can not be quantified, our findings suggest that the "hype" about such changes is likely overblown.[1]

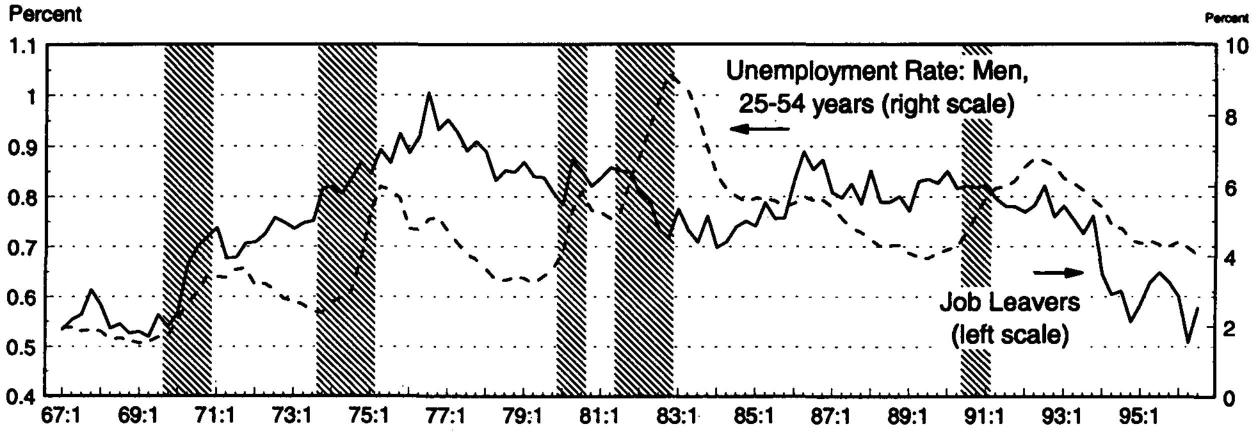
---

1 We also extended our investigation beyond a consideration of only the individual predicitve content of the labor market variables for compensation growth. Specifically, we augmented the set of regressors in equation (2) to include the average duration of unemployment, job losers, jobleavers, involuntary part-time workers, and the job apprehension index, but we were unable to reject the null hypothesis that their coefficients are jointly equal to zero. In addition, we used principal components analysis and extracted the first principal component from the above set of variables. However, when the first principal component was included as a regressor in equation (2), it displayed a statistically and quantitatively insignificant effect.

Chart 17

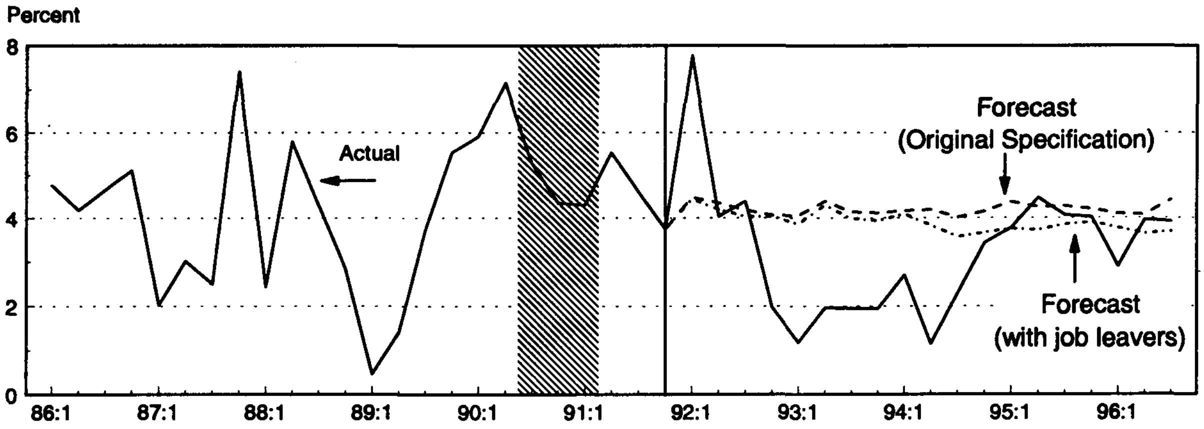
Job leavers as a percent of civilian labor force and the unemployment rate

1967Q1-1996Q3



Compensation growth out-of-sample forecast

1992Q1-1996Q3



Note: Shading refers to NBER recessions.

Conclusion

In this article we explored the issue of an inflation puzzle and the lack of acceleration in price inflation during the current expansion. In particular, we investigated whether the appearance of an inflation puzzle reflected a shift in the inflation process or could be linked to factors or developments emanating from the labor market. Drawing upon the Phillips-curve literature, we specified a model for the rate of change in prices and found little evidence of instability in the estimated relationship over the post-1991 period.

The analysis then examined the key determinants of price inflation and uncovered evidence suggesting that unusual restraint in unit labor costs – principally through the behavior of compensation – has been central to the continued low inflation environment observed over the current expansion. Paralleling the exercise for price inflation, we estimated a Phillips-curve model for compensation growth and documented the marked weakness in this variable during the period 1992-94.

In an attempt to understand the source(s) of the compensation growth shortfall, we examined several hypotheses concerning possible changes in labor market behavior. Our results provided little support for such explanations. Aggregate data series representing the decline in the number of union workers, the increased number of contingent workers, the increased average duration of unemployment, the decline in job security, and the level of consumer confidence concerning job prospects all failed to provide much of an account for the compensation growth shortfall. Only one series – job leavers as a percent of the civilian labor force – seemed to play a role in the compensation growth shortfall, although its explanatory content was limited and quite modest. Thus, it now appears that it is the behavior of compensation growth and its dramatic slowdown during the early 1990s that remain a topic for debate and an area for future research.

Appendix: Derivation of the accelerationist Phillips-curve model

This Appendix provides a brief discussion of the accelerationist model of the Phillips curve from equations (1) and (2). The key feature of this version can be illustrated by examining the relationship between the output gap (and unemployment gap with a constant NAIRU) and the inflation rate. Abstracting from the influence of other terms, note that the system of equations (1) and (2) can be rewritten as:

$$INF_t = \alpha_1 GDPGAP_{t-1} + \sum \alpha_{2+i}\left(INF_{t-i}\right) + \sum \alpha_{7+i}\left(LXNG_{t-i}\right) \quad (3)$$

and

$$LXNG_t = \left(\beta_3 U_{t-1} + \sum \beta_{3+i} INF_{t-i}\right) \Big/ \left(1 - \beta_1 L - \beta_2 L^2\right) \quad (4)$$

where we use the definition of unit labor cost growth as compensation growth less productivity growth in equation (3) and L denotes the lag operator in equation (4) such that $L^k X_t = X_{t-k}$.

We can substitute equation (4) into equation (3) to obtain an expression relating current inflation to the output gap, the unemployment gap, and past rates of inflation. If the sum of the coefficients on lagged inflation equals unity, then there is a "natural rate" value of the output gap (and unemployment gap) of zero that is consistent with a constant rate of inflation. Alternatively, this model would associate a permanent positive value for the output gap with an ever-accelerating inflation rate. Within our system of equations, the condition that the sum of the coefficients on lagged inflation equals unity is given by:

$$\alpha_3 + \alpha_4 + \alpha_5 + \left[\left(\alpha_8 + \alpha_9\right)\left(\beta_4 + \beta_5\right) / \left(1 - \beta_1 - \beta_2\right)\right] = 1. \quad (5)$$

The hypothesis that the coefficients on lagged inflation sum to unity can be tested using the OLS estimates of equations (1) and (2) to construct an estimate of the expression on the left-hand side of equation (5) and its standard error. The standard error is the standard error of a function of several estimated parameters and can be computed using the delta method approximation (Greene (1993), p. 297):

$$SE\left[g(\theta)\right] = \sqrt{\left(\partial g / \partial \theta'\right) VAR(\theta) \left(\partial g / \partial \theta\right)}$$

where θ denotes the parameters in equation (5), $g(\theta)$ is the function of the parameters in (5), and

VAR(θ) is the variance-covariance matrix of those parameters.

Because of the slight disparity in the sample periods for Table 1 and Table 2, equations (1) and (2) were both estimated over the period 1967Q2-1996Q3. The estimate for the expression on the left-hand side of equation (5) was 0.87 with an estimated standard error of 0.08. Thus, we are unable to reject the null hypothesis that the sum of the coefficients in (5) is equal to unity at the 5% significance level.

Data appendix

Inflation equation:

INF: Growth Rate of Core CPI, all urban consumers. Monthly, SA, 1982-84=100.
Source: Department of Labor, Bureau of Labor Statistics, *Consumer Price Index.*

UNITG: Growth in Unit Labor Costs, Nonfarm Business Sector. Quarterly, SA, 1992=100.
Source: US Department of Labor, Bureau of Labor Statistics, "Productivity and Costs".

GDPGAP: Logarithmic ratio of GDP-to-POTGDP, where GDP is quarterly real gross domestic product and POTGDP is potential GDP, quarterly. Both variables are in 1987 dollars until 1987Q3 and in chain-weighted 1992 dollars from 1987Q4-present.
GDP source: National Income and Product Accounts. Potential GDP source: FRBNY Staff estimate.

$OILG^*$: Net positive change in the real price of oil. The percentage change in the current real price of oil from the previous year's maximum if positive and zero otherwise.
Source: Data for the price of oil are an extension of Mork's (1989) series which reflect corrections for the effects of price controls during the 1970s. The real price of oil is constructed by deflating the nominal price index by the GDP deflator.

Compensation equation:

LXNG: Growth rate of Compensation per Hour, Nonfarm Business.[2] Quarterly, SA, 1992=100.
Source: US Department of Labor, Bureau of Labor Statistics, "Productivity and Costs".

INF: Growth Rate of Consumer Price Index for all items, urban consumers. Monthly, SA, 1982-84=100.
Source: Department of Labor, Bureau of Labor Statistics, *Consumer Price Index.*

U: Unemployment Rate for Male, 25-54 years. Monthly, SA.
Source: Department of Labor, Bureau of Labor Statistics, "The Employment Situation", Table A-8.

UIR: Unemployment insurance per job loser, normalized by the average annual earning of a manufacturing worker – a replacement ratio which shows what fraction of earnings (of manufacturing workers since they are the ones most likely to collect unemployment insurance) is replaced by unemployment insurance = (*YPTU/LUJL*)/(*YPWF/LAMANU*), where:

> *YPTU*: Government unemployment insurance benefits. Quarterly, SAAR.
> Source: National Income and Product Account, *Gross Domestic Product*, Table 2.1.

---

2 Wages and salaries for workers plus employers' contributions for social insurance and private benefit plans. Except for nonfinancial corporations, the series also includes an estimate of wages, salaries, and supplemental payments for the self-employed.

LUJL: Job losers and persons who completed temporary jobs. Monthly, SA.
Source: Department of Labor, Bureau of Labor Statistics, "The Employment Situation", Table A-6.

YPWF: Wage and salary disbursements in Manufacturing. Quarterly, SAAR.
Source: National Income and Product Account, *Personal Income*, Table 2.1.

LAMANU: Nonfarm Payroll Employees in Manufacturing. Monthly, NSA.
Source: Department of Labor, Bureau of Labor Statistics, "The Employment Situation", Table B-1.

SOC: Quarterly.
Source: FRBNY Staff Estimate.

DUM: accounts for the restraining effect of the wage and price freeze in 1971Q4 and the rebound after the relaxing of controls in 1972Q1 (=1 in 1971Q4, -0.6 in 1972Q1, and 0 elsewhere).

Additional variables examined:

Commodity Price Index: Spot commodity price index for all commodities. Monthly, 1967=100.
Source: *K.R. CRB Commodity Index Report*, Knight-Ridder Financial.

Trade-Weighted Value of the Dollar: United States vs. eighteen countries, based on 1984 bilateral trade weights for the eighteen currencies. Monthly, 1980=100.
Source: Federal Reserve Bank of Atlanta.

Number of Union Members: Wage and salary employees who are members of unions. Also includes members of employee associations which are similar to labor unions. Unbenchmarked and available on an annual basis.
Source: Department of Labor, Bureau of Labor Statistics, *Employment and Earnings.* (January) Table 58.

Number of Work Stoppages: Work stoppages (beginning in period) involving 1000 or more workers. Monthly, NSA.
Source: US Department of Labor, Bureau of Labor Statistics, *Compensation and Working Conditions.*

Job Leavers: Number of civilians unemployed who left their job. Monthly, SA.
Source: Department of Labor, Bureau of Labor Statistics, "The Employment Situation", Table A-6.

Average (Mean) Duration of Unemployment: Monthly, SA.
Source: Department of Labor, Bureau of Labor Statistics, "The Employment Situation", Table A-5.

Temporary Workers: Temporary employees on nonfarm payroll. Monthly, NSA. Source: Department of Labor. Bureau of Labor Statistics. "The Employment Situation", Table B-1.

Involuntary Part-Time Workers: Part time workers for economic reasons (frequently referred to as "partially unemployed"), non-agricultural industries. Excludes workers who usually work full time but worked part time (1 to 34 hours) during the reference weeks for holidays, illness and inclement weather. Monthly, SA.
Source: Department of Labor, Bureau of Labor Statistics, "The Employment Situation", Table A-3.

Consumer Apprehension about Job Loss: Constructed variable: negative of present employment situation component of the Consumer Confidence Index (= 2.5 (*CCIN* - 0.6**CCIEN*), where *CCIN*: Consumer confidence, 1985=100 and *CCIEN*: Consumer expectations, 1985=100).
Source: The Conference Board.

Bibliography

Blanchard, Olivier J. (1984): "The Lucas Critique and the Volcker Deflation". *American Economic Review Papers and Proceedings* 74, pp. 211-215.

Blomberg, S. Brock and Ethan S. Harris (1995): "The Commodity-Consumer Price Connection: Fact or Fable?" Federal Reserve Bank of New York, *Economic Policy Review*, October, pp. 21-38.

Boldin, Michael D. (1995): "Problems in Determining the Capacity-Inflation Link". *Mimeo*, Federal Reserve Bank of New York.

Bradsher, Keith (1995): "Greenspan Says Fears Over Jobs Have Held Down Wage Growth". New York Times, October 20, p. D8.

Brown, R.L., J. Durbin and J.M. Evans (1975): "Techniques for Testing the Constancy of Regression relationships Over Time". *Journal of the Royal Statistical Society*, No. 37, pp. 149-192.

Chow, Gregory C. (1960): "Tests of Equality Between Sets of Coefficients in Two Linear Regressions". *Econometrica*, No. 28, pp. 591-605.

Eggert, Robert (ed.): "Blue Chip Economic Indicators". Published by Capitol Publications, Inc. Alexandria Va. 1991-95.

Englander, Steven A. and Cornelis A. Los (1983): "The Stability of the Phillips Curve and Its Implications for the 1980s". Federal Reserve Bank of New York, *Research Paper*, No. 8303, February.

Friedman, Milton (1968): "The Role of Monetary Policy". *American Economic Review*, No. 58, pp. 1-17.

Fuhrer, Jeff (1995): "The Phillips Curve is Alive and Well". *New England Economic Review*, March/April, pp. 41-56.

Goodfriend, Marvin (1993): "Interest Rate Policy and the Inflation Scare Problem: 1979-1992". Federal Reserve Bank of Richmond, *Quarterly Review*, Winter, pp. 1-24.

Gordon, Robert J. (1970): "The Recent Acceleration of Inflation and its Lessons for the Future". *Brookings Papers on Economic Activity*, Vol. 1, pp. 8-41.

Gordon, Robert J. (1975): "The Impact of Aggregate Demand on Prices". *Brookings Papers on Economic Activity*, Vol. 6, pp. 613-62.

Gordon, Robert J. (1977): "Can the Inflation of the 1970s Be Explained?" *Brookings Papers on Economic Activity*, Vol. 8, pp. 253-77.

Gordon, Robert J. (1982): "Inflation, Flexible Exchange Rates, and the Natural Rate of Unemployment", in M. N. Baily (ed.), *Workers, Jobs and Inflation*. Washington: Brookings Institution, pp. 88-152.

Gordon, Robert J. (1990): "US Inflation, Labor's Share, and the Natural Rate of Unemployment", in Heinz König (ed.), *Economics of Wage Determination*. Berlin and New York: Springer Verlag, pp. 1-34.

Gordon, Robert J. (1996): "The Time-Varying NAIRU and its Implications for Economic Policy". *NBER Working Paper*, No. 5735, August.

Greene, William H. (1993): *Econometric Analysis*. New York: Macmillan Publishing Company.

Hamilton, James D. (1996): "This is What Happened to the Oil Price-Macroeconomy Relationship". *Journal of Monetary Economics*, No. 38, pp. 215-20.

King, Robert G. and Mark W. Watson (1994): "The Post-War US Phillips Curve: A Revisionist Econometric History". *Carnegie-Rochester Conference Series on Public Policy*, No. 41, pp. 157-219.

King, Robert G., James H. Stock and Mark W. Watson (1995): "Temporal Instability of the Unemployment-Inflation Relationship". Federal Reserve Bank of Chicago, *Economic Perspectives*, May/June, pp. 2-12.

Macroeconomic Advisers, LLC (1996): "The US Economic Outlook", October 8.

Meyer, Laurence H. (1997): "The Economic Outlook and Challenges for Monetary Policy". Remarks at the Charlotte Economics Club, January 16.

Mork, Knut A. (1989): "Oil and the Macroeconomy When Prices Go Up and Down: An Extension of Hamilton's Results". *Journal of Political Economy*, No. 97, pp. 740-44.

Passell, Peter (1995): "Has This War Been Won?" *New York Times*, October 25, p. D1.

Phelps, Edmund S. (1967): "Phillips Curves, Expectations of Inflation and Optimal Unemployment Over Time". *Economica*, No. 34, pp. 254-81.

Phillips, A.W. (1958): "The Relation between Unemployment and the Rate of Change of Money Wage Rates in the United Kingdom, 1861-1957". *Economica*, No. 25, pp. 283-300.

Sullivan, Daniel (1996): "Unemployment Duration and Labor Market Tightness". *Federal Reserve Bank of Chicago Letter*, March.

Spiers, Joseph (1995): "Why Inflation Will Keep Falling". *Fortune Magazine*, October 2, p. 59.

Spiers, Joseph (1996): "It's Fear of Inflation That's Spooking the Market". *Fortune Magazine*, June, p. 22.

Tootell, Geoffrey M.B. (1994): "Restructuring, the NAIRU, and the Phillips Curve". *New England Economic Review*, September/October, pp. 31-44.

Valletta, Robert (1996): "Has Job Security in the US Declined?" *Federal Reserve Bank of San Francisco Weekly Letter*, No. 96-07, February 16.

Comments on: "A look at the US inflation puzzle" by Cara S. Lown and Robert W. Rich

by Gabriele Galati

This paper addresses the question of whether there is an inflation puzzle, in the sense that core CPI inflation has been lower since 1991 than unemployment and capacity utilisation figures appear to tell us. It is an interesting paper that comes up with a story behind the observed behaviour of these variables and follows a straightforward estimation strategy to test it. The story is that the inflation dynamics have not changed during the recent expansion, and the only fundamental determinant of inflation that behaved differently is compensation growth. On the other hand, there is no evidence of recent changes in labour market conditions. The authors conclude therefore that there is no evidence of an inflation puzzle and that low compensation growth in the early 90s is the major factor behind the low level of inflation.

To provide evidence for their story, the authors follow a strategy based on three steps. First, they estimate an equation that describes a price-inflation Phillips curve (equation (1) in the paper) which differs from standard specifications in that it includes unit labour costs among the explanatory variables. The authors then perform structural break tests and dynamic out-of-sample forecasts to investigate its stability. They find a reasonably good fit and no evidence of structural breaks and conclude that the inflation dynamics haven't changed recently.

It may be useful to start off with estimating a more "traditional" price-inflation Phillips curve that does not include unit labour costs as a right-hand side variable and examine whether the inflation dynamics described by this equation have changed recently. It might well be that this equation exhibits a structural break around 1991. I would then add the variable measuring unit labour costs to the right-hand side variables and report that this modified price-inflation equation is stable over the whole period. These two results together could then be used to argue that compensation growth may be a factor explaining the recent lack of acceleration of inflation as opposed to other factors such as monetary policy having become more credible.

The paper also needs some more motivation of the choice of explanatory variables. It seems that the authors include variables on the basis of the significance of their coefficient. For example, they include an oil shock variable that measures only positive real oil price changes because a variable measuring negative oil price changes turned out not to have a significant coefficient. There is some discussion in Section 1 of fundamental factors that drive inflation and what the authors call "shock factors". This distinction enters the price-inflation equation, for example with an oil shock variable that is meant to capture exogenous supply shocks, but not in a systematic way. It may be useful to report how the inclusion of variables that capture the other shock factors that are mentioned, for example commodity prices or the exchange rate, affects the results.

The authors then examine whether the behaviour of the determinants of price inflation (the right-hand side variables in equation (1)) has changed since 1992. Graphical evidence suggests that unit labour costs have exhibited an unusually weak growth rate after 1991, and that this has been caused mainly by weak compensation growth. To test more formally for a change in the determination of compensation growth, the authors estimate a compensation growth Phillips curve (equation (2) in the paper) which explains hourly compensation growth in terms of the lagged unemployment rate, lagged inflation and other variables capturing labour market conditions. They then subject this equation to the same stability tests performed on the price inflation Phillips curve. Although different tests fail to provide evidence of structural breaks, the out-of-sample forecasts consistently overpredict actual compensation growth between 1992 and 1994. The authors suggest that recent changes in the labour market may therefore explain the low inflation observed during the current expansion.

To test this hypothesis, the authors first carry out an informal, graphical analysis of the behaviour of different indicators of labour market conditions. Once they find an indicator that has behaved differently over the last few years, they use it in the compensation growth equation as an explanatory variable. Since none of the indicators has a significant coefficient or improves the fit and out-of-sample forecasting ability of the equation, they conclude that there is no empirical evidence supporting the hypothesis.

One problem with this strategy is that compensation growth (the left-hand side variable in equation (2)) and labour market conditions (a right-hand side variable) are endogenous and jointly determined. It is therefore difficult to infer any causal relation from labour market conditions to compensation growth based on single coefficient estimates or measures of performance of the whole equation. Some labour market variables may be relevant even if their coefficients are not significant in the compensation growth equation. One way to solve this problem – although difficult to follow in practice – is to look for truly exogenous variables that may have affected labour market conditions over the recent years and then test for their significance in the compensation growth equation. Examples include changes in the laws on subsidies for hiring unemployed, changes in tax breaks, etc.

Another problem is how to measure the effect of labour market variables that matter only in last part of the sample period. The authors solve this problem by adding the variable to the equation and verifying its coefficient over the whole period but admit that this makes it difficult to capture appropriately the effect of a change in labour market conditions. An alternative testing strategy would be to use a more flexible specification and introduce the exogenous variables measuring changes in labour markets both as dummy variables and as interaction term with other explanatory variables. The impact of these on the fit and forecasting ability of the compensation growth equation may lead to different conclusions on the role of changing labour market conditions.

Expectations, learning and the costs of disinflation
Experiments using the FRB/US model

Antúlio Bomfim, Robert Tetlow, Peter von zur Muehlen and John Williams[1]

Introduction

In the first quarter of 1980, inflation in the United States was over 16% and over the ten years ending in 1982Q4, the average rate of inflation rate was 8.8%.[2] The outbreak of inflation in the 1970s in the United States and elsewhere engendered a greater concern among central bankers worldwide of the cost of inflation, and with it, the employment costs of disinflation. Over the period from 1983Q1 to 1996Q3 – that is, after the Volcker disinflation – US inflation averaged 3.5%.

In many respects, the Volcker disinflation was unique in the American experience. It coincided with changes in Fed operating procedures, most notably in the adoption of monetary targeting. The disinflation also commenced from a point at or near a local maximum in the inflation rate. Perhaps more importantly, the ex post sacrifice ratio of one is relatively low by historical standards.[3] According to Ball (1994), the sacrifice ratio for this episode was about one; that is, it required the equivalent of a one-percentage-point increase in the unemployment rate for one year for each percentage point reduction of inflation.[4] Table 1 documents much of the history of disinflations in the US and shows the Volcker disinflation as the least costly in recent history. The most notable features of Table 1 are, first, that disinflations have always been costly in terms of foregone employment, and, second, that the cost has varied significantly with the case and with the method of measurement.

These observations raise the issue of whether the way in which a disinflation is carried out significantly affects the costs that must be borne along the transition path. More generally, the question for policy makers is "what are the sources of costly inflation adjustment and is there a way in which monetary policy can affect these costs?"

In 1996, a group of economists at the Federal Reserve Board completed the first working version of a new macroeconometric model of the US economy. The product of several years of work, version 1.0 of FRB/US replaced the venerable MPS model for model-based forecasting and policy analysis at the Fed in the spring. But as the name change hints, FRB/US is more than just a refurbishing of its predecessor. The builders of FRB/US sought to exploit the many changes in the

---

1 The views expressed in this article are those of the authors and are not necessarily those the Board of Governors of the Federal Reserve System or its staff. The authors thank, without implication, Flint Brayton and Dave Reifshneider for helpful remarks and Steve Sumner for able research assistance.

2 Inflation, in this context, means the rate of change from a year earlier of the quarterly consumer price index, (all items, all urban areas).

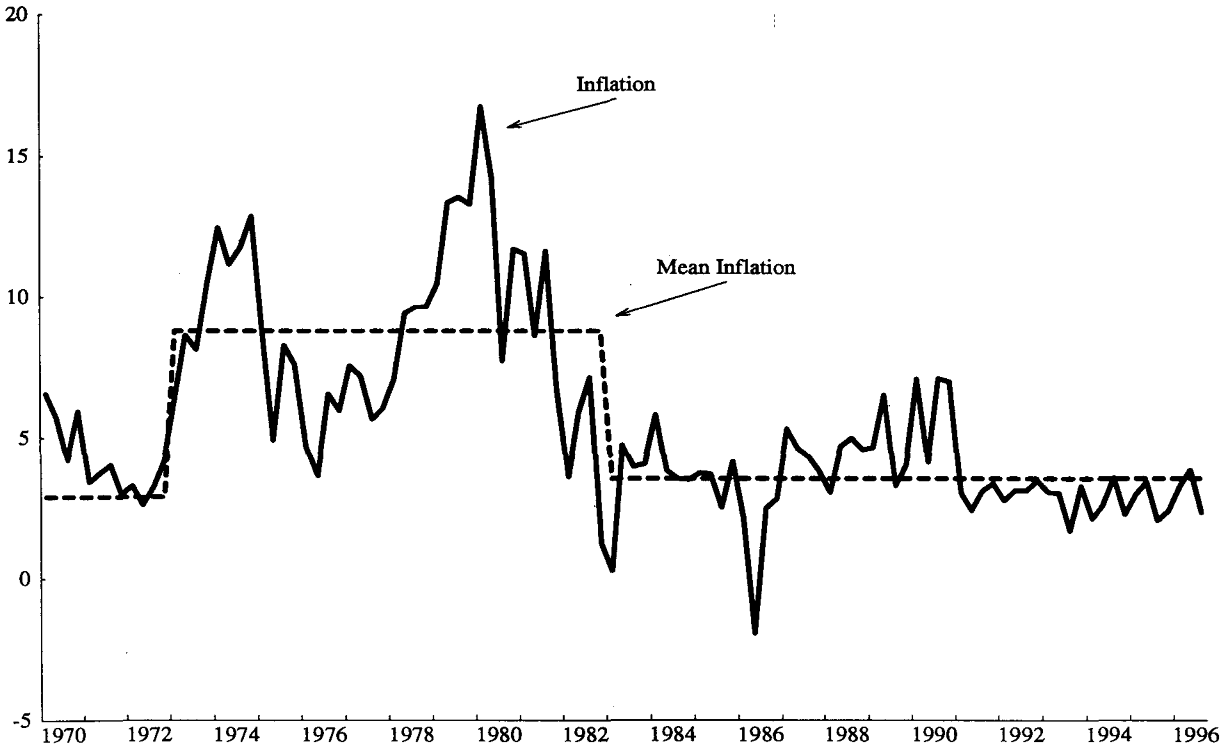
3 Henceforth, except as noted, the term sacrifice ratio shall be taken as being the undiscounted employment cost of a one percentage point reduction of inflation. In the simulations that we present, the versions of the model used are either linear or nearly so and therefore the magnitude of the disinflation is of no consequence. Similarly, since the model's ex post Okun coefficient is very nearly the same regardless of the way in which a disinflation is carried out, there is very little incremental information to be had from examining output sacrifices.

4 Ball actually computed an output sacrifice ratio of 1.8 which can be converted to an employment sacrifice ratio of about one using standard measures of the Okun coefficient of about 1½ to 2.

economic modelling and econometrics that have taken place over the twenty-five years since MPS was first put into service. The wish list for the model was a long one. The new model had to fulfill a role in both forecasting and policy analysis. To serve as a tool of policy analysis, the model had to be able to characterize outcomes as arising from the conscious decisions of agents. In short, the agents represented by the model's behavioral equations had to be optimizing, albeit subject to constraints. The model had to produce results to well-posed policy questions that are sensible from the point of view of general equilibrium or systems properties. At the same time, goodness of fit, assessed on an equation-by-equation basis, remained a priority. The adoption of recent innovations in modeling technology has served to improve the tradeoffs that are inherent in reconciling these often conflicting objectives.

Figure 1

Quarterly growth, at an annual rate, of the consumer price index (all urban, all items)



Disinflation involves the conscious act of a monetary authority to change the prevailing expectations of private agents. Unlike the bulk of earlier models, the FRB/US model can offer some meaningful insight on policy issues of this sort. Most macroeconometric models are not able to address the issue of changing expectations because their dynamic lag structures do not distinguish between sources of propagation that are attributable to *intrinsic dynamics*, meaning adjustment costs, irreversibilities, vertical input-output relationships and the like, and *expectational dynamics*, stemming from errors in expectations formation and learning. In this way, these models fall prey to the Lucas critique.

The FRB/US model addresses the Lucas critique:[5] model dynamics stem from a confluence of intrinsic and expectational propagation mechanisms. Intrinsic adjustment costs are

[5] We say that FRB/US "addresses" rather than solves the Lucas critique in that the model lays out behavioral decision rules for private agents. This allows users to experiment with the expectational responses to policy interventions, as we do in this paper. However, since private decisions are not modeled simultaneously with and conditioned upon

modeled using a polynomial adjustment costs (PAC) technology described by Tinsley (1993) and Brayton and Tinsley (1995). PAC technology avoids the empirical rejection of rational expectations models of optimizing behavior that have plagued the literature. This is done through a straightforward generalization from the more traditional level adjustment cost (LAC) specification, in which costs are quadratic in the *level* of the decision variable, to allow high-order adjustment costs.[6] An important advance with the new model is that it can be simulated using either model-consistent expectations, or expectations that are based on limited information, in this case generated from a vector autoregression. It is this flexibility in the modeling of expectations that allows us to investigate the importance of expectations formation for the cost of bringing down inflation within a working, econometric macro model. The model also allows the use of learning rules to come to grips with the transition dynamics surrounding fundamental changes in the economic environment. Following Sargent (1993), one might argue that "rational" expectations may be a meaningful characterization of an economy at its stochastic steady state, but it may be important to consider "nonrational" transition dynamics stemming from the learning of the new environment associated with a change in policy target.

Table 1

Some estimates of sacrifice ratio for the United States

Reference	Period	Method	Sacrifice ratio	
			Output	Employment
Okun (1978)	various	econometric survey	6 - 18	2 - 6
Gordon (1982)	1920 - 1921	calculation	0.2	-
	1929 - 1933	calculation	1.7	-
Gordon-King (1982)	1954 - 1980	econometric	4.3 - 5.8	-
Sachs (1985)	1981 - 1984	calculation	3.0	1.7
	1984 - 1987	output simulation	3.7[1]	-
Blinder (1987)	1982 - 1985	calculation	-	2.1
Mankiw (1991)	1982 - 1985	calculation	2.8	1.9
Ball (1994)	1969 - 1971	calculation	2.9	1.5[2]
	1974 - 1976	calculation	2.4	1.2[2]
	1980 - 1983	calculation	1.8	0.9[2]
DRI (1995)[3]	1967 - 1994	simulation	-	2.7
MPS (1995)[4]	1963 - 1993	simulation	3.6	1.8

Note: Okun's Law calculations used to convert from output sacrifice ratios to employment sacrifices are those of the reference author, except where otherwise stated.

[1] Simulation figure refers to computations based on Sachs (1983) Table 8, p. 172. [2] Converted from output to employment using an Okun's Law of 2 as suggested by footnote 5, p. 170. [3] Based on a simulation of the July 1995 version of the DRI Quarterly Model of the US Economy. [4] Unpublished results from a simulation conducted on the 1995 version of the MPS model.

The rest of this paper proceeds as follows. We begin, in Section 1, with a brief summary of the FRB/US model, focusing on those aspects of the model that are pertinent to the policy issue at hand. This naturally leads us to discuss wage and price determination and the characterization of monetary policy. Section 2 outlines our approach to learning. Section 3 considers disinflation experiments under VAR-based and model consistent expectations, with and without learning. In those

behavioral decision rules for the monetary and fiscal authorities, it would be an overstatement to say that the current version of the model solves the Lucas critique.

6 Or adjustment costs in the moving-average of the level of the variable. It turns out that these are equivalent although sometimes the resulting estimates suggest one interpretation over the other. On this, see Tinsley (1993).

instances where we study learning under model-consistent expectations it is advantageous for numerical reasons for us to linearize the model.[7] We compare our results with those in the literature and with the estimated benefits of inflation reduction. The final section offers some concluding remarks and discusses directions for future research.

1. The FRB/US model

This section reviews the structure and methodological underpinnings of FRB/US with a view towards providing an understanding of the sources of costly price and inflation adjustment. To facilitate this end, our discussion of the model will be brief and concentrated primarily on the wage-price block. The model's wage-price block is a bit more complicated than a single-equation specification for inflation and so in order to show the value-added from this modest increase in complexity, we shall compare the model with standard Phillips curves. We begin, however, with the broader aspects of the model's construction.

1.1 General aspects

In comparison with most working econometric models, including MPS, FRB/US is small, with about 300 equations overall, of which about 50 are behavioral. A large part of the reason for the smaller size is the higher emphasis placed on system properties in model evaluation. More so than in the past, the success or failure of FRB/US hinges on its ability to answer meaningful policy questions with a minimum of fuss. This means that behavior is more often modeled from a "top-down" perspective than a "bottom-up" one. About half of these behavioral equations are explicitly derived as decision rules governing the behavior of representative agents acting with foresight to achieve explicit objectives in the presence of constraints.

One of these constraints is the information set to which agents have access in forming their expectations. As we shall see below, the model can be simulated with any of a number of different characterizations of the availability of information and its subsequent gathering over time. For the purposes of estimation, however, it was assumed that information is constrained to a small set of aggregate macroeconomic variables, summarized by a low-order vector autoregression (VAR) described below. This low-order VAR forms the core of the expectations generation system that is augmented by one or two other variables, the precise nature of which varies from agent to agent and from sector to sector.

Wrapped around this expectations structure are the decision rules which lay out the planned paths for the decision variables under control of each agent, given expectations and the costs of adjustment. In the case of financial variables, these intrinsic costs of adjustment are regarded as negligible; equilibrium is determined by arbitrage.[8] However for real variables, the presence of adjustment costs means that agents must balance the cost of being away from their desired level for a variable with the cost of adjusting to get to that desired level. This compels them to plan ahead, subject to expected future conditions, to set out a path for the adjustment toward the desired level. An important step forward in the empirical performance of macroeconomic models based on decision rules is adoption of polynomial adjustment costs (PAC) in place of the traditional level-adjustment-cost (LAC) decision rules pioneered by Hansen and Sargent (1980). Relaxing the restriction that

---

[7] Little is lost from linearization because the full model is quite close to being linear itself.

[8] With the exception of the monetary policy rules, which have the form of interest-rate reaction functions, we shall have nothing further to say about the modeling of financial variables. Interested readers are invited to consult Kozicki et al. (1996) or Brayton and Tinsley (1996) for details.

adjustment is costly only in the level of the decision variable dramatically reduces the problem of excess residual autocorrelation that has traditionally plagued empirical tests of macroeconomic decision rules.

And at its heart, FRB/US is a neoclassical growth model. That is, once all the dynamic adjustment is complete, the model settles down on a balanced growth with consumption determined by wealth, wealth determined in large part by productivity and steady-state real interest rates, and real interest rates determined by the savings-investment decisions of private agents in concert with government. The model is non-Ricardian. Of the several reasons which account for this feature, the most important is that risk-averse agents "overdiscount" their future labor income, in the same spirit as the Blanchard (1985) model of perpetual youth. More generally, fiscal policy can act upon the long-run equilibrium of the economy through purchases, transfers, government debt, and distortionary taxes. Developments in the understanding of the econometrics of stochastically trended variables play a substantial role in the determination of the characteristics of the long-run equilibrium of the model. But since we take as given that to be useful for policy analysis a model must contain within its structure a well-defined steady state, this requirement has, upon occasion, overridden the results of cointegration tests.

1.2 Modeling wages and prices

Inflation dynamics in the FRB/US model are driven by adjustment costs in nominal prices and wages and, for some specifications, by the characterization of expectations and learning. Let us set aside expectations for the moment so as to focus on the structural determinants of costly disinflation in the model. We have already noted that the existence of costly adjustment obliges agents to plan ahead to achieve their objectives for some date in the future. The target level of the variable in question may be moving over time, which in turn means that agents lay out a path for the target itself as well as a plan to close the gap between current levels of wages (or prices) and the target level.

Taking, for the moment, the target level of the producer price as given, the job of firms as price setters is to choose the producer price, p, to minimize:

$$L=\sum_{i=0}^{\infty}\beta^{i}\left[b_0\left(p_{t-i}-p^{*}_{t-i}\right)^2+b_1\left(p_{t-i}-p_{t-i-1}\right)^2+\sum_{k}b_{k+1}\left(\Delta^{k}p_{t-i}-\Delta^{k}p_{t-i-1}\right)^2\right] \tag{1}$$

where p^* is the target price, $b = [b_0, b_1,..., b_{k+1}]$ is a vector of parameters measuring the intrinsic costs of adjustment as well as the cost of being away from the target price, and β is the subjective rate of discounting.[9] Equation (1) is more general than the standard level adjustment cost (LAC) problem in that costs might be borne in adjusting in higher-order changes in prices, as determined by b_2 and k.[10] Just as Nickell (1985) has shown for the LAC problem, the Euler equation derived from this problem can be solved to arrive at an error-correction representation of the following form:

$$\Delta p=-A(1)\left(p_{t-1}-p^{*}_{t-1}\right)+\sum_{j=1}^{k}\gamma_j(b)\Delta p_{t-1}+\sum_{i=0}^{\infty}\rho_i(\beta,b,k)E_{t-1}\Delta p^{*}_{t+i} \tag{2}$$

---

9 In this equation and all others, all variables other than those naturally expressed as rates, are expressed in natural logarithms, unless otherwise noted. Variables expressed as rates include interest rates and inflation.

10 Tinsley (1993) shows that extending the quadratic adjustment cost model either to allow quadratic costs of higher-derivative changes, as in equation (1), or to moving averages of the level of the decision variable, can produce the same Euler equation. Which interpretation is most useful will vary from case to case, often depending on the specific empirical results.

where $A(L) = 1 + a_1(b)L + a_2(b)L^2 + ... + a_{k+1}(b)L^{k+1}$. Equation (2) differs from what one would derive from the LAC model in two significant ways: first, the LAC model implies that lags in Δp do not appear in the decision rule; second, under LAC the forward term in which the Δp^*_{t+i} appears is constrained to be a geometrically declining lead. Under PAC technology, both of these restrictions are relaxed. It is worth emphasizing, however, that notwithstanding the relaxation of the constraints noted, important testable restrictions across parameters still exist.

We shall briefly outline form the empirical results for our price equation, but before doing so, we need to discuss the specification of target variables, p^*_{t+i}, and expectations thereof.

It is useful to dichotomize the determinants of the target price level into stationary and non-stationary components. The specification of the stationary components is essential for satisfying long-run equilibrium conditions: prices equal marginal cost (plus a fixed markup) in the long run and the producer real wage equals the marginal product of labor. The stationary components of target variables impart dynamics on the model, in this case, a cyclical markup.

The first-order conditions for profit maximization, given Cobb-Douglas production technology in three factors, imposes the necessary long-run restriction on wages and prices, given productivity and the price of energy.[11] This determines the nonstationary component of the target price:

$$\hat{p}_t = \delta + \phi(w_t - \lambda_t) + (1 - \phi)p^e_t + \varepsilon_t \qquad (3)$$

where w is the wage rate, λ is trend labor productivity and p^e is the price of energy.

Prices are assumed to be set by firms and wages to be set or negotiated jointly by firms and workers such that equation (3) holds as an equilibrium implication of the combined behavior of wages and prices. One way to think of the bargaining process that might bring this about is that the markup of prices over costs (and thus the real wage) moves with excess supply of labor in reflection of the ebb and flow of bargaining power over the business cycle. Adding this stationary component to $\hat{p}_t$ and taking expectations yields the target price:

$$p^*_{t+i} = E_{t-1}\left[\hat{p}_{t+i} + \phi^p U_{t+i}\right] \qquad (4)$$

where $\hat{p}$ is the solution to equation (3) above, ignoring the error term, and $\phi^p U_{t+i}$ is the cyclical portion of firms' target markup of prices over costs. A target wage rate equation also exists, modeled in an analogous fashion, and contains its own cyclical determinants. These two equations jointly satisfy equation (3) and, in steady state, settle down on the equilibrium real wage.

The estimated PAC wage and price equations indicate that there is stickiness in rates of change and in the acceleration of each. From the estimated equations, the cost parameters can be computed. In the case of the price equation, we have $(b_0, b_1, b_2) = (0, 99.1, -17.5)$. The zero coefficient on the LAC term is a homogeneity restriction; that is, $b_0 = 0$ ensures that it is not costly for firms to adjust prices at the steady-state rate which implies the wage-price block is consistent with any targeted inflation rate. The dominant feature of the PAC technology for price-setting firms is the large coefficient penalizing changes in the inflation rate; inflation is very sticky in FRB/US. Interestingly, $b_2 < 0$ means that while firms find changes in inflation costly, when inflation must change, firms prefer that it take place rapidly. These results contrast with pure price stickiness models such as Taylor (1980), Calvo (1983) and Rotemberg (1983) since, without inflation stickiness, these models imply that costless disinflation is feasible. Finally, we observe that the estimates also show that price setting behavior carries a weight of 43% on forward elements and 57% on predetermined elements.

---

11 The production function is Cobb-Douglas with constant returns to scale in three factors, capital, labor and energy. A constant term picks up the equilibrium level of price markups over cost.

Table 2

Some estimates of the producer price equation
1963Q1-1994Q4

$\Delta p = 0.10\left(p_{t-1} - p^{*}_{t-1}\right) + 0.38\Delta p_{t-1} + 0.19\Delta p_{t-2}$	SEE = 0.0025 $R^2 = 0.88$
$+0.38\Delta\left(\frac{\tilde{p}^{e}}{p}\right) - 0.05\Delta\left(\frac{\tilde{p}^{e}}{p}\right)_{t-1} + 0.43\sum_{i=0}^{\infty}\tilde{\rho}_i(A,\beta,b,k)\Delta p^{*}_{t+i}$	BG(1) = 0.71 J(20) = 0.47
$p^{*}_{t} = 0.98(w - \lambda) + 0.02p^{e} - 0.003u$	

Notes: variable definitions: p is the price of output excluding government, farm and energy; p^* is the desired price; p^e is the price of energy consumption; $\Delta(\tilde{p}^e / p)$ is the change in the scaled relative price of energy consumption; $w$is the nominal wage rate measured by the employment cost index; λ is the level of trend labor productivity for the production sector defined above; and u is the level of the demographically adjusted unemployment rate. Some constant terms have been suppressed. BG(n) is the probability value associated the Breusch-Godfrey test of serial correlation for up to n lags. J(k) is the probability value associated with Hansen's test of overidentifying restrictions.

Most of the empirical specifications of PAC equations include extra terms designed to capture idiosyncratic elements of the particular sector's behavior. The price and wage equations are no different. The producer price equation contains a term (at two dates) that captures the faster pass-through into producer prices of energy prices than other costs. Table 2 provides the highlights of the empirical results for the price equation. The structure of the wage equation is similar to that of prices and so we shall not explicitly review the empirical specification here. It is worth mentioning, however, that in the wage equation the extra terms are more numerous; they include, a dummy for the Nixon wage-and-price controls, a relative minimum wage variable, and a variable capturing the implications of payroll taxes. Also, as it turns out, three lags of the change in the wage rate enter the wage equation and wages turn out to be stickier than prices.

Costly inflation adjustment, as we have here, is also a feature of the real wage contracting model of Buiter-Jewitt (1981) and Fuhrer-Moore (1995), and of standard accelerationist Phillips curves. It implies that disinflation will be costly in terms of forgone output regardless of expectations. That is, while expectations formation in general and the credibility of policy in particular may impinge on the results of disinflation experiments, they will not do so in such as way as to reduce the costs of disinflation to zero. Since the standard accelerationist Phillips curve specifies that inflation is costly to adjust, but does not provide an avenue for policy to affect these costs, the FRB/US wage-price specification encompasses the accelerationist Phillips curve model and extends it by allowing both forward and backward-looking elements to influence inflation and by disentangling the intrinsic and expectational sources of propagation.

1.3 Expectations formation

In laying out equation (4), we did not discuss how expectations of future target prices are formed. In use for forecasting, and during the course of estimation, it is assumed that expectations formation can be represented by a low-order vector autoregression.[12] Effectively, what this does is to restrict the full, reduced-form of the model-consistent solution to one of a smaller order that might

[12] The use of VAR-generated expectations obviates the need to use simultaneous equations methods such as FIML or three-stage least squares which, while feasible in principle, turn out to be too unwieldy to pursue in practice for a model the size of FRB/US.

arguably be more in keeping with what might be used by agents facing costly information collection and processing. Thus, we think of the VAR-generated expectations solution of the model as being a reasonable approximation to the limited-information solution of the model.

The precise VAR used varies from sector to sector but three core equations are used in all cases. These core VAR variables are the output gap, the inflation rate, and the nominal federal funds rate. This core VAR represents the aggregate information that is available to all agents in the economy. On top of this, agents in each sector have at their disposal local information that is useful for forecasting variables of concern to them. Localized agents take macro-aggregate variables as given, and so the structure of these local VARS is restricted such that the macro-aggregate variables of the core VAR affect the local variables, but not vice versa.

The specification of the core VAR differs from traditional VARs in two small but important ways. The structure is summarized in equations (5) and (6) below. First, there are restrictions imposed on the $A(\circ)$ matrix to ensure that the VAR is consistent with any target level of inflation. The second and more important difference – related to the first – is that the long-run levels of the variables are constrained by an endpoint condition:

$$A(L)x = u_t \tag{5}$$

$$x = \begin{bmatrix} y - y^{\infty} \\ \pi - \pi^{\infty} \\ r - r^{\infty} \end{bmatrix} \quad u = \begin{bmatrix} u^{y} \\ u^{\pi} \\ u^{r} \end{bmatrix} \tag{6}$$

where y is the deviation of output from potential output, measured in percent; π is, once gain, inflation; r is the nominal federal funds rate; and the ∞ superscript indicates the transversality or endpoint condition. With a traditional (stationary) VAR, the sample mean serves as the "endpoint" for the system; that is, the point upon which forecasts will eventually settle. In the core VAR used in FRB/US, this is generalized to permit systematic changes in long-run conditions. Since FRB/US is a natural rate model, the endpoint constraint for the output gap is trivially identical to zero. For inflation and the federal funds rate, however, the endpoint constraint represents the private sector's long-run expectation for inflation and, through the imposition of the Fisher identity, the real interest rate, respectively. In the data, the endpoint for the (nominal) federal funds rate is computed from the implied forward rates at the far end of long-term bond rates, less a term premium which is taken as constant in the long run. The endpoint for inflation is taken from a survey of inflation expected over a 10-year period kept by the Federal Reserve Bank of Philadelphia.[13]

The presence of endpoint variables in the VAR reflects the finding of nonstationarity in nominal interest rates in the historical data. This could represent permanent or near-permanent shifts in the real interest rate. Alternatively, it could represent either shifts in the (perhaps implicit) targeted inflation rate, or the inability of the Fed to achieve a fixed (implicit) target due to unfavorable shocks. The data suggest that both real rate movements and inflation movements have been at work.

The important point to take from the inclusion of these endpoint constraints is that they can move endogenously. In the absence of moving endpoints, one is faced either with interest rates that are "too smooth" in in-sample simulation, relative to the historical data, or "too volatile",

---

[13] More precisely, it is only for the more recent period that the inflation endpoint is taken from the Philadelphia survey. There are significant gaps in the Philadelphia survey series and it starts only in 1980Q4. To fill in the missing dates and extend the series a little bit, the Philadelphia survey was spliced with the now-discontinued Hoey survey series. For dates in common, the two are very similar. But even the Hoey survey does not go back before 1979. For periods prior to then, we use an econometrically constructed series described in Kozicki and Tinsley (1996).

depending on whether one models nominal interest rates as stationary or integrated.[14] By itself, this modeling strategy is a topic worthy of a lengthy discussion that we cannot provide here; interested readers can find the pertinent arguments in Kozicki et al. (1996).

The question at hand is what governs the movements of these endpoints. On this, we shall focus on the movements of the endpoint for inflation and since inflation is a monetary phenomenon, it stands to reason that the endpoint for inflation is determined by monetary policy. This is not to say that the endpoint is the long-run target for monetary policy. Rather, it is the role of the monetary authority to elicit changes in private agents' expectations so as to guide expected inflation in the long run, π^{∞}, to the target inflation rate: π^*. Notice, then, that we have implicitly divided expectations formation into two parts, distinguished by frequency: we have expectations over the short run given, in some cases at least, by the VAR $(x \mid x^{\infty})$, and we have lower frequency dynamics, governed by movements in π^{∞}. This presents a myriad of hypothetical experiments regarding the response of private agents to policy interventions. In particular, as we shall demonstrate in the next section, VAR-based expectations can be replaced with model consistent expectations. And both systems of expectations formation allow instantaneous recognition of a change in the policy target, or learning. We regard this flexibility as one of the major contributions of the model.

Table 3

Estimates of the core VAR

1963Q1 - 1994Q4

$\pi = 0.62\pi_{t-1} + 0.03\pi_{t-2} + 0.30\pi_{t-3} - 0.11\pi_{t-4} + 0.16\pi^{\infty}$ $+0.16r_{t-1} - 0.26r_{t-2} + 0.01r_{t-3} + 0.07r_{t-4} + 0.01r^{\infty}$ $+0.07y_{t-1} - 0.03y_{t-2} + 0.05y_{t-3} + 0.03y_{t-4}$	SEE = 1.13 $R^2 = 0.26$
$y = -0.02\pi_{t-1} + 0.03\pi_{t-2} - 0.04\pi_{t-3} - 0.03\pi_{t-4} + 0.03\pi^{\infty}$ $-0.01r_{t-1} - 0.40r_{t-2} + 0.29r_{t-3} - 0.05r_{t-4} + 0.18r^{\infty}$ $+1.09y_{t-1} - 0.07y_{t-2} - 0.02y_{t-3} - 0.05y_{t-4}$	SEE = 1.12 $R^2 = 0.33$
$r = 0.14\pi_{t-1} + 0.00\pi_{t-2} + 0.05\pi_{t-3} - 0.12\pi_{t-4} - 0.06\pi^{\infty}$ $+0.99r_{t-1} - 0.44r_{t-2} + 0.49r_{t-3} - 0.11r_{t-4} + 0.07r^{\infty}$ $+0.31y_{t-1} - 0.13y_{t-2} - 0.10y_{t-3} + 0.02y_{t-4}$	SEE = 1.14 $R^2 = 0.30$

Notes: The specification shown is equivalent to the representation of equations (5) and (6). Variable definitions: r is the nominal federal funds rate measured on an effective basis; π is the quarterly change in the chain-weighted personal consumption price index, measured at annual rates; y is the deviation of output from its trend, measured in logarithms. r^{∞}, $\pi^{\infty}$are defined in the text.

1.4 Monetary policy

Consistent with the handling of expectations, monetary policy in the model is also flexible. The most coherent characterization is that of monetary policy being governed by explicit feedback rules that have as their medium-to-long-term objective the establishment and maintenance of a nominal anchor. It is under these circumstances that the consistency of private expectations and policy intentions can be addressed.

[14] The similarity of this issue to that of the "excess volatility" or "excess smoothness" of consumption for a given data generating process for income is not a coincidence. The same issues of identifying the "size of the unit root" and its source are at work here.

The monetary policy rule that has been used in our assessments of model properties is the very same equation that is used in VAR to model expectations. In this instance, the core VAR's federal funds rate equation is more than simply a linearized, restricted-information solution to the larger system but is literally the policy rule. When the model is simulated in this mode, it is assumed that the process of learning the authority's rule and target is already complete. This completes the chain of logic on the estimation strategy that was followed: VAR expectations are used to identify the expectational dynamics of the model, PAC technology then selects the intrinsic dynamics, and the VAR-based policy rule validates the expectations that private agents were assumed to have used in the first place. For the purposes of policy analysis, however, one can replace the VAR-based policy rule with another rule and examine the process of the adaptation of private expectations to a "data generating process" for policy that is different from that of history.

The VAR rule represents the average historical experience over the period from 1963 to 1994. We are more interested in specific episodes and in specific rules, and so in our simulations below, we pay scant attention to the VAR rule. Instead, we consider two alternative rules which both approximate monetary policy in the United States over the period during and after the Volcker disinflation. One is the well-known Taylor (1993, 1994) rule. The other carries the same arguments as the Taylor rule but allows for a richer dynamic structure; we refer to this rule as the Post-1970s rule.

The Taylor rule is a very simple federal funds rate reaction function, loosely fitted to the experience of the 1980s. It calls upon the Fed to respond to contemporaneous deviations of inflation from its target level and to deviations of (the log of) output from its potential level. Both behavioral arguments in the reaction function carry the same 0.5 coefficient. We have replaced the contemporaneous value of inflation in Taylor's specification with a four-quarter moving average. Otherwise, the rule we use is as Taylor has written it.

Table 4

Two federal funds rate reaction functions

$r = rr^* + \sum_{i=0}^{3}\pi_{t-i}/4 + 0.5\left(\sum_{i=0}^{3}\pi_{t-i}/4 - \pi^T\right) + 0.5y_t + \mu_t$	Taylor Rule (calibrated)
$r = 0.62r_{t-1} - 0.06r_{t-2} + 0.25r_{t-3} + 0.07r_{t-4}$ $+0.09(\pi - \pi^T) + 0.17(\pi_{t-1} - \pi^T) + 0.07(\pi_{t-2} - \pi^T)$ $+0.23(\pi_{t-3} - \pi^T) + 0.84y_t - 0.21y_{t-1} - 0.10y_{t-2} - 0.27y_{t-3}$ $+(1 - 0.62 + 0.06 - 0.25 - 0.07)\left[rr^* + \sum_{i=0}^{3}\pi_{t-i}/4\right] + \mu_t$	Post-1970s rule (1979Q4-1994Q4) SEE = 1.17 R^2 = 0.93

Notes: The specification shown is equivalent to the representation of equations (5) and (6). Variable definitions: r is the nominal federal funds rate measured on an effective basis; π is the quarterly change in the chain-weighted personal consumption price index, measured at annual rates; y is the deviation of output from its trend, measured in logarithms. r^∞, π^∞ are defined in the text.

The Taylor rule has two features that some observers might take as being unusual. The first is that the endogenous variable is written in levels instead of changes. A rule written in levels requires the authority (and the modeler) to have a good estimate of the steady-state real interest rate, shown here as rr^*. In principle, a rule written in Δr could be allowed to error-correct to an unknown rr^*. However, arriving at a rule with good properties would oblige the researcher to begin with a good idea of rr^*so the differences between good rules written in Δr and r are probably more apparent than real. The second noteworthy feature of the Taylor rule is that there are (effectively) no dynamics in the equation. This issue we can address with our Post-1970s rule. The two rules are shown in Table 4.

The lag structure of the Post-1970s rule implies that the response of the federal funds rate to deviations of target variables from their desired levels builds over time. The other important difference between the Taylor and Post-1970s rules is that the latter is a much more forceful rule. The steady-state coefficient on inflation in the Post-1970s rule is about 3, or six times as large as that of the Taylor rule. Similarly, the steady-state weight on the output gap is about 2, or four times as large as the Taylor rule.

1.5 Learning

As we have already discussed, expectations formation in the model can be dichotomized into parts: expectations over the short to medium term, and expectations of the long-term; that is, of the endpoints. We can illustrate differences in the method of forming expectations in the short and medium term through our choice of model-consistent expectations, or VAR-based expectations. We can also investigate expectations formation at a lower frequency by modeling expectations of the endpoint for inflation. Table 5 illustrates the connection between our dichotomization of expectations and the literature. We think of issues such as "credibility" and "reputation" as having to do with knowledge of, belief in, and acceptance of, the target rate of inflation of the monetary authority; this is a low-frequency concept that can be made distinct from the dynamic resolution of expectations errors along the transition path to the new target inflation rate.[15] Expectational errors in the short run – apart from those attributing to incorrect perceptions of the target – may come from limited information, as we have when we use VAR-based expectations, or from unanticipated shocks.

Table 5

A mapping between modelling of expectations in FRB/US and the venacular

Short- to medium-term⇨		**Dynamic expectations formation**	
long-term		VAR-based	Model consistent
endpoint expectations	discontinuous	limited information but credible	"rational" and credible (perfect foresight)
	continuous	limited information and learning	"rational" but not credible; learning

Note: By "discontinuous", we mean that the endpoint for inflation is jumps instantaneously to the new target value; "continuous" means that the expected target moves endogenously and continuously in response to observable information; that is, that agents learn adaptively.

The usefulness of this distinction is best seen by considering the example of a pre-announced disinflation. If the public believes the Fed is sincere in its announcement, and is capable of carrying out the policy, the announcement can be said to be "credible".[16] In this case, the endpoint for inflation jumps discontinuously to the new, announced level. If private agents also form expectations using the entire structure of the model, then expectations are model consistent. Full credibility combined with model consistency and pre-announcement gives perfect foresight. This case is

[15] The distinction between expectations of the long run for inflation and the short run is often blurred in academic treatments of the subject which typically use models with no dynamic features other than credibility.

[16] See Blackburn and Christensen (1985) for a good survey of credibility of monetary policy.

represented by the north-east quadrant of the body of Table 5. There is, however, no reason to assume that because agents find an announcement fully credible that they are also have access to full information, in which case agents are using limited information – VAR-based in our cases – to summarize the economy and the north-west quadrant applies. The lower part of the table refers to cases of learning. In these cases, agents use observable information to update their expectations of the true target rate of inflation.

There is a substantial literature documenting a number of ways one could model endogenous learning of the target rate of inflation. Bullard (1991) provides a good, short survey at an elementary level with Sargent (1993) offering a more in-depth treatment, which is still readable by nonspecialists. In our treatment, we assume that agents know with certainty the rule that is being used to conduct monetary policy but treat the target rate of inflation as being a random variable, the behavior of which may be reasonably approximated by a random walk. Under these circumstances, regardless of the particular rule, there are two source of uncertainty emanating from monetary policy: a transitory shock to the federal funds rate itself, and a once-and-for-all shock to the target. Agents must infer the target rate of inflation by solving a signal extraction problem. Let us take the Taylor rule, which we repeat for convenience, as an example:

$$r = rr^* + 1.5\left(\sum_{i=0}^{3} \pi_{t-i} \Big/ 4\right) + 0.5y_t - 0.5\pi_t^T + \mu_t \tag{7}$$

Agents observe Δr and need to estimate π^T from their observations. Changes in the federal funds rate are given by:

$$\Delta r = f(observables) - 0.5\Delta\pi^T + \mu_t + \mu_{t-1} \tag{8}$$

The actual law of motion for the target rate of inflation is:

$$\pi_t^T = \pi_{t-1}^T + v_t \tag{9}$$

Using this, let the surprise part of the change in the federal funds rate be:

$$\lambda_t = \mu_t - \mu_{t-1} - 0.5v_t \tag{10}$$

Only the left-hand side of equation (10) is observable. Assuming that both μ and v are independently and identically distributed, the solution to the signal extraction problem is for agents to use the average occurrence of the shocks, as measured by the signal-to-noise ratio of the two disturbances, as their tool for forecasting the target inflation rate:

$$E_t v_t = \left[\frac{0.5^{-2}\sigma_v^2}{0.5^{-2}\sigma_v^2 + 4\sigma_\mu^2}\right]\lambda_t = k\lambda_t \tag{11}$$

This means, in the parlance of Kalman filtering, that agents will update their perceptions of the inflation target by the Kalman gain, k:

$$\pi_t^* = \pi_{t-1}^* + k\lambda_t \tag{12}$$

In principle, one could estimate the model of expectations updating and allow k to be time-varying. If, for example, a monetary authority had been practicing a policy of allowing the inflation rate to drift, as our learning model says, and then reformed itself to a fixed target, then $\lim_{t\to\infty} \sigma_v^2 = 0$ and $\lim_{t\to\infty} k = 0$, and the data generating process for the target would eventually be perceived as being stationary. Our efforts in this area are ongoing but success has been elusive to date. Thus, for the experiments of the next section, we will be experimenting with a variety of constant gain specifications.

2. The costs of disinflation

We are interested in the unemployment costs of a permanent 1% disinflation.[17] We compute the sacrifice ratio under a variety of different conditions. To provide benchmarks, it is useful to begin by setting aside learning and conducting one disinflation experiment with each of the full-information and limited-information versions of the model, under the assumption of full credibility. Recall that under these circumstances, we reduce the target inflation rate by one percentage point and the perceived target drops simultaneously. For this benchmark case, we use the Taylor rule. The results are best summarized graphically as in Figure 2 below.

Perhaps the most striking thing about Figure 2 is the similarity of the two simulations. When monetary policy is fully credible, the dynamic movements of key variables are quite comparable. There are some noteworthy differences, but they are relatively minor: full credibility permits the nominal federal funds rate to fall immediately in the model-consistent case, but not so in the limited-information case, there is a small low-frequency secondary cycle in the limited-information case, and inflation begins its descent a bit later in the limited information case. The overall similarity of the results says two things about the model: first, that granting the monetary authority full credibility is giving agents a lot of information, and second, that the VAR-representation of expectations encompasses quite well the information agents need to forecast future events with reasonable accuracy, for this shock. Disinflation is initiated by a disturbance to the federal funds rate, a variable that appears directly in core VAR. The disinflationary impetus – beyond what credibility alone provides – operates through the unemployment rate, and the output gap, which moves in parallel with unemployment for shocks from the demand side of the model, appears in the core VAR. Finally, inflation itself also appears in the core VAR. Given this information, it is perhaps not entirely surprising that the two disinflations are broadly similar.

What of our sacrifice ratios? The sacrifice ratio for the model-consistent expectations simulation ends up being about 1.6 percentage points. For the limited-information simulation, the corresponding number is 1.4.[18] Referring back to Table 1, a benchmark figure for the sacrifice ratio (in unemployment space) might be 2 to 2.5. Hence, we are a bit low, but not by a large amount.

Now let us consider model behavior under learning. We show, in Table 6, a sampling of output of four different values for the updating parameter, *k*, at three different dates: five years after the beginning of the intervention, ten years after and twenty. The top panel summarizes the results using the model-consistent expectations version of the model[19] while the bottom panel shows the results for the limited-information version. In both cases, it is the Taylor rule that governs monetary policy. The numbers in the table are basis points of change in absolute terms so that "1.00" means that full convergence has been achieved.

Whereas in Figure 2 we saw remarkable similarity, here we see interesting differences. The bottom line of each panel in the table shows the results for the simulations shown in Figure 2. Let us begin by contrasting the sacrifice ratios for those simulations, with the sacrifice ratios for the

---

[17] The model is very nearly linear so the magnitude of the disinflation is unimportant. What is more important is the "permanence" since one cannot reasonably argue that the costs of disinflation can be considered if the lower inflation rate cannot be sustained without incurring further costs. This means that the structural stability of the model is crucial to the results.

[18] The fact that disinflation under full credibility with limited information is less costly than under full information will be surprising to those who recall the opposite results from the new classical literature of the 1970s and early 1980s. Here, however, the errors in expectations along the transition path evidently aid the process of disinflation. The precise source of this phenomenon, which is fairly robust to the choice of policy rule, is under investigation.

[19] For the model-consistent expectations simulations under learning we use a linearized version of the model. This avoids a great many numerical problems and saves time at only a small cost since the model is close to linear.

slowest rates of learning, 2.5% per quarter. In the case of model-consistent expectations, the difference in sacrifice ratios from full credibility and very slow learning rate is only two-fold: a sacrifice ratio of 3.2 after 20 years when agents learn slowly, versus 1.6 when they immediately believe a disinflation is forthcoming. In the limited-information case, however, the differences are more marked; the slow-learning case has a sacrifice ratio over six times larger than the fully credible case.

Figure 2

FRB/US disinflation experiments (deviations from base case)
Limited information versus model-consistent expectations
Instantaneous recognition of policy change from federal funds rate surprises

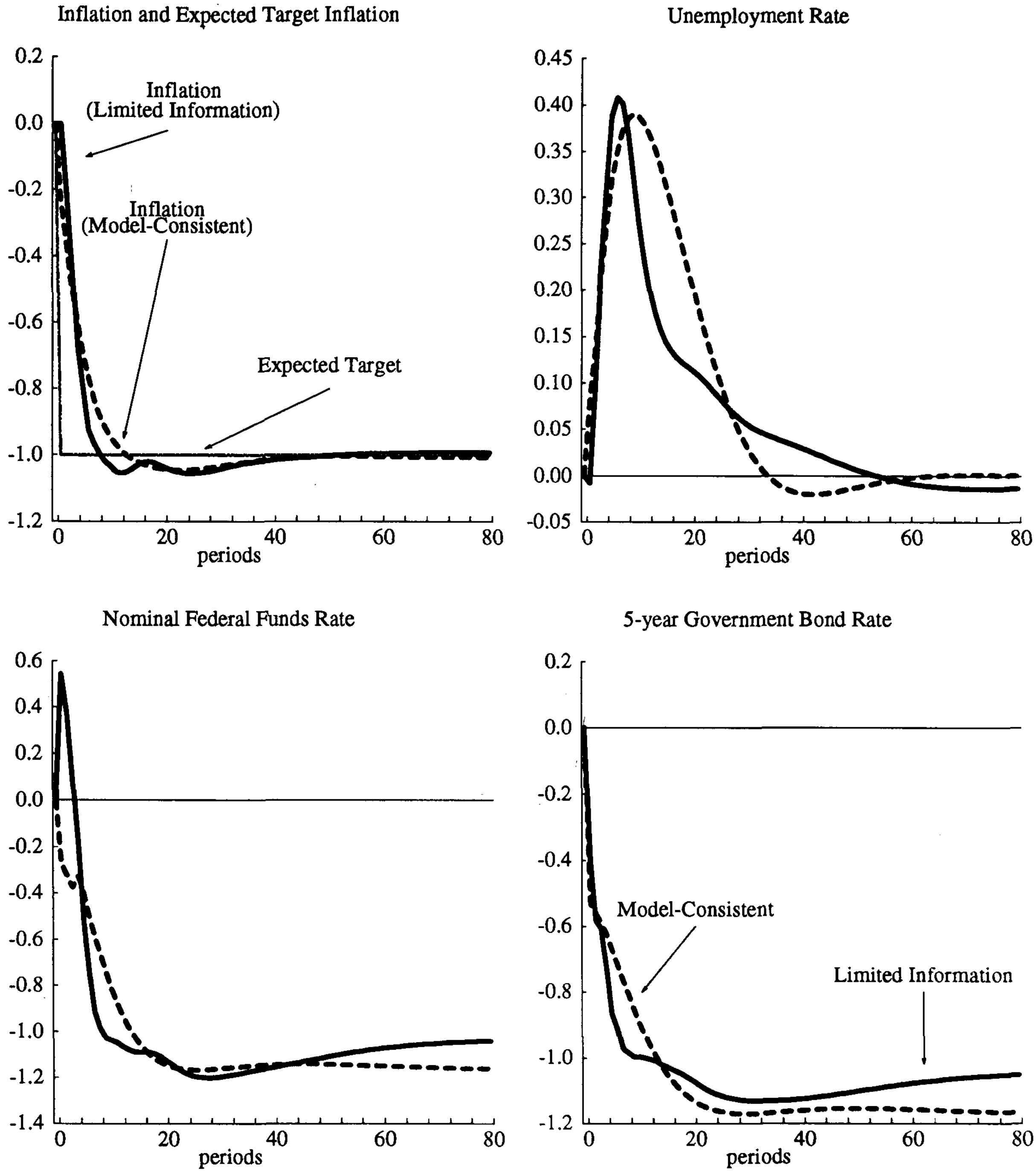


Table 6

FRB/US model disinflation experiments
Model-consistent expectations versus VAR-based expectations
Learning from federal funds rate surprises with the Taylor rule

Model-consistent expectations (at selected years)									
Learning rate	**Expected target**			**Inflation**			**Sacrifice ratio**		
%	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
2.5	0.38	0.63	0.86	0.42	0.74	0.95	2.0	2.6	3.2
5	0.62	0.86	0.98	0.56	0.88	0.99	1.7	2.2	2.5
10	0.86	0.98	1.00	0.72	0.96	0.97	1.5	1.9	2.0
25	0.99	1.00	1.00	0.83	0.95	0.95	1.5	1.7	1.7
Credible	1.00	1.00	1.00	0.96	1.01	1.00	1.4	1.7	1.6
VAR-based expectations (at selected years)									
Learning rate	**Expected target**			**Inflation**			**Sacrifice ratio**		
%	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
2.5	0.07	0.14	0.30	0.32	0.40	0.51	2.6	5.1	8.6
5	0.14	0.27	0.50	0.37	0.49	0.65	2.3	4.2	6.3
10	0.25	0.47	0.77	0.44	0.63	0.84	2.0	3.2	4.3
25	0.48	0.80	1.03	0.60	0.87	1.02	1.5	2.3	2.7
Credible	1.00	1.00	1.00	1.03	1.01	0.99	1.0	1.4	1.4

Note: Sacrifices are measured in terms of cumulative increase in unemployment, in percentage points, divided by the change in inflation relative to the baseline as of the date indicated.

One reason for this difference is that in the absence of credibility, monetary policy in the limited-information case must do more of its work through the standard Keynesian channel of creating excess supply. We see this in the fact that actual inflation runs consistently ahead of the perceived target. A large part of the reason for this is that longer-term bond rates, which are modeled using the expectations theorem of the term structure do not move in advance of movements of aggregate demand in the absence of credibility. This implies higher sacrifice ratios, and a front-loading of the costs of disinflation. Slow learning implies slow adjustment of inflation in the limited-information case; even at the remarkably fast rate of 25% per quarter, the expected target is only 80% of the way to the authority's objective after ten years.

In the model-consistent case, the expected target inflation rate and the actual rate are much closer together. Interestingly, only at the slowest speed of learning does the expect target rate of inflation lag behind actual inflation after five years. And the economy gets a lot of benefit for a relatively small amount of learning; a 10% learning rate produces sacrifice ratios that are quite close to the full-credibility figures, and drives the expected target to inside of 15% of the way to its destination within five years.

There is a school of thought that suggests that a "cold shower" disinflation will be less costly in terms of forgone employment by eliciting a discrete jump in the perceived target inflation rate at the outset of the exercise. Running against this argument is the standard Keynesian notion that if price and wage rigidities are structural, then policy that fails to take into consideration the pace at which they can adjust will end up creating more and longer lasting unemployment than is necessary for the task. This issue is examined in Table 7 where we compare the performance of the Post-1970s policy rule with the Taylor rule. In order to give expectations the best possible chance to jump rapidly to a new, lower level we use model-consistent expectations. The upper panel of the table is identical to Table 6.

Table 7 shows that indeed the harsh medicine of the Post-1970s rule delivers a more rapid decline in the perceived target rate of inflation than the Taylor rule, for all rates of learning. However this speedier disinflation is at the cost of larger accumulated employment losses. Since even the fully credible disinflation is more costly than with the Taylor rule, evidently the Post-1970s rule operates against binding constraints in terms of adjustment costs. In essence, the Post-1970s rule cannot produce federal funds rate surprises at any faster a rate than it produces excess supply.[20] There are, however, other characterizations of monetary policy that may support cold shower disinflations. For example, if one were to model central bankers choosing, unobserved by private agents, high or low inflation regimes based on their "conservativeness", then it might be possible for conservative central bankers to signal their intentions by choosing a tougher course of disinflation than a less conservative central banker would choose to mimic.

Table 7

FRB/US model disinflation experiments
Taylor rule versus the Post-1970s rule under model-consistent expectations
Learning from federal funds rate surprises

Taylor rule (at selected years)									
Learning rate	**Expected target**			**Inflation**			**Sacrifice ratio**		
%	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
2.5	0.38	0.63	0.86	0.42	0.74	0.95	2.0	2.6	3.2
5	0.62	0.86	0.98	0.56	0.88	0.99	1.7	2.2	2.5
10	0.86	0.98	1.00	0.72	0.96	0.97	1.5	1.9	2.0
25	0.99	1.00	1.00	0.83	0.95	0.95	1.5	1.7	1.7
Credible	1.00	1.00	1.00	0.96	1.01	1.00	1.4	1.7	1.6
Post-1970s rule (at selected years)									
Learning rate	**Expected target**			**Inflation**			**Sacrifice ratio**		
%	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
2.5	0.63	0.78	0.92	0.80	0.90	0.97	2.9	4.5	5.7
5	0.86	0.95	0.99	0.94	1.00	0.99	2.4	3.6	4.1
10	0.98	1.00	1.00	1.01	1.00	0.99	2.1	3.0	3.1
25	1.00	1.01	1.00	1.00	1.00	0.99	2.0	2.5	2.5
Credible	1.00	1.00	1.00	0.98	1.00	0.98	1.9	2.2	2.2

Note: Sacrifices are measured in terms of cumulative foregone employment divided by the change in inflation relative to the baseline as of the date indicated.

We can now compare, in broad terms, the sacrifice ratios of Table 6 and Table 7 with our stylized facts from Table 1. If we take as before, 2 to 2½ as benchmark numbers for sacrifice ratios,

[20] One way in which the monetary authority could bring about a larger initial surprise and reduce the perceived target would be to add a large one-time discretionary jump in the federal funds rate at the start of the disinflation. This may be what some people mean when they speak of "cold shower" disinflations. Doing so, however, goes against the notion of monetary policy being governed by rules and violates the assumed orthogonality of the temporary and permanent shocks to the policy rule.

we observe that one needs either slow learning with otherwise rational expectations, or fast learning with limited information to reconcile the model with the historical sacrifice ratio. A calibration of a learning rule, using survey data, to the Volcker disinflation can yield a learning rate anywhere between 2% and about 15%, depending on the precise specification of the rule and on the particular survey, with perhaps the most trustworthy numbers being in the 5% range. These observations suggest it is hard to reconcile the VAR-expectations version of the model under smooth learning rules at about a 5% rate with the benchmark sacrifice ratio. One possibility is that learning has been more discrete than our constant-gain specification permits. This notion has some intuitive appeal since, although the errors are steadily diminishing, in the Taylor rule simulations with limited information, agents learning at a rate of 5% per quarter are still making errors of the same sign even twenty years after the initiation of the disinflation. Given the aforementioned preferences of price setters for discrete jumps in inflation when prices must move at all, one might expect "second-derivative" learning as well as the first derivative type modeled here.[21]

Table 8

FRB/US model disinflation experiments
Taylor rule versus the Post-1970s rule under VAR-based expectations
Learning from federal funds rate surprises

Taylor rule (at selected years)									
Learning rate	**Expected target**			**Inflation**			**Sacrifice ratio**		
%	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
2.5	0.07	0.14	0.30	0.32	0.40	0.51	2.6	5.1	8.6
5	0.14	0.27	0.50	0.37	0.49	0.65	2.3	4.2	6.3
10	0.25	0.47	0.77	0.44	0.63	0.84	2.0	3.2	4.3
25	0.48	0.80	1.03	0.60	0.87	1.02	1.5	2.3	2.7
Credible	1.00	1.00	1.00	1.03	1.01	0.99	1.0	1.4	1.4
Post-1970s rule (at selected years)									
Learning rate	**Expected target**			**Inflation**			**Sacrifice ratio**		
%	**5**	**10**	**20**	**5**	**10**	**20**	**5**	**10**	**20**
2.5	0.14	0.25	0.46	0.60	0.54	0.68	2.0	5.2	8.0
5	0.26	0.45	0.72	0.70	0.66	0.84	1.8	4.2	5.7
10	0.46	0.72	0.94	0.85	0.81	0.97	1.6	3.2	3.9
25	0.72	1.08	1.00	1.13	1.08	1.00	1.4	2.0	2.4
Credible	1.00	1.00	1.00	1.07	0.98	1.00	1.3	1.5	1.4

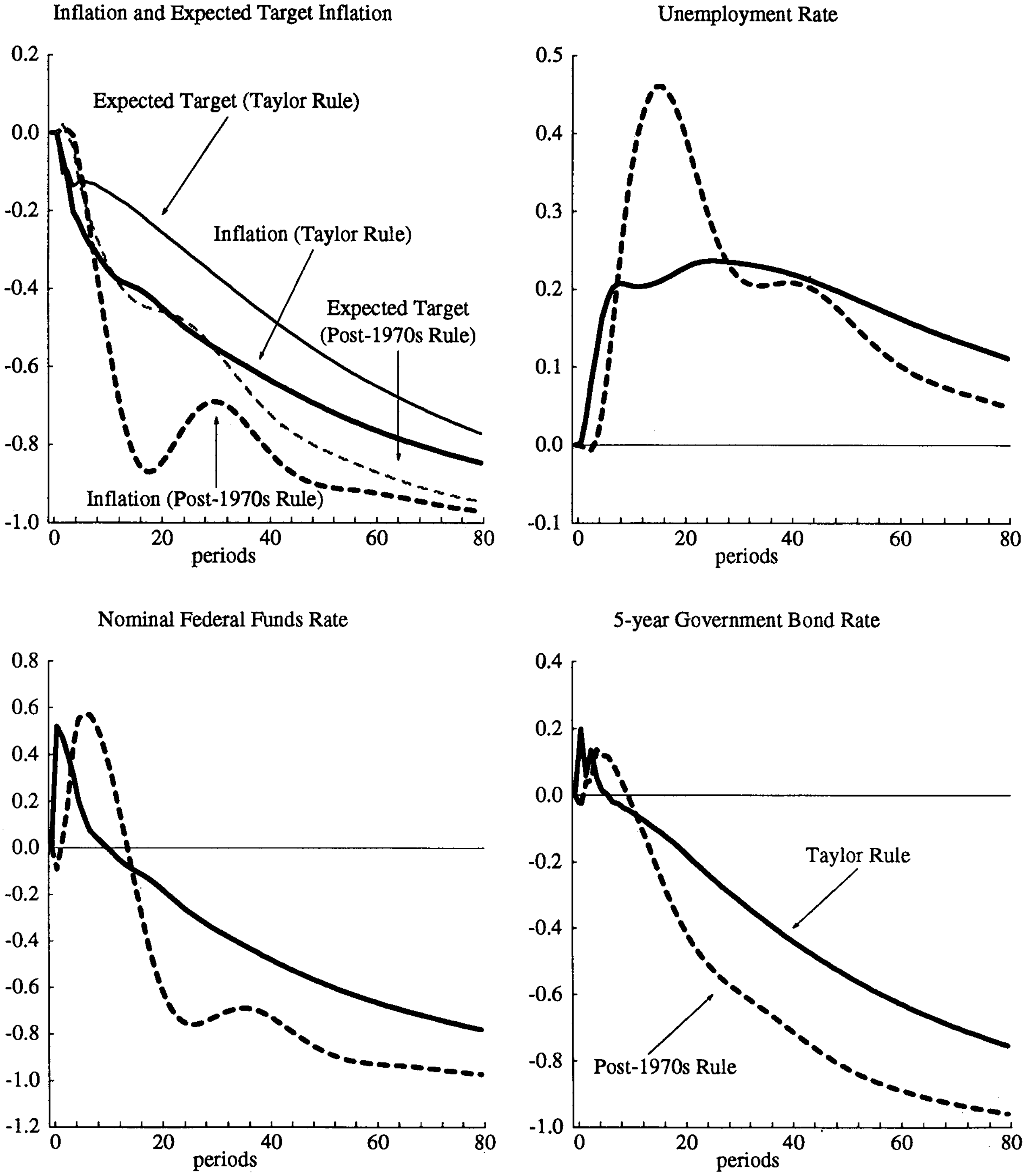
Note: Sacrifices are measured in terms of cumulative foregone employment divided by the change in inflation relative to the baseline as of the date indicated.

---

[21] We must also acknowledge the fact that many, if not most, of the sacrifice ratios in Table 1 are not computed to correspond with so well-defined an experiment as the ones we have considered for this paper. In the case of Ball (1994), for example, sacrifice ratios were computed between periods when inflation was constant for an arbitrary period of time on the accelerationist logic that if inflation is constant, the output gap must be closed (or equivalently, the unemployment rate is at the NAIRU). In a stochastic economy, however, a zero gap is neither necessary nor sufficient for inflation to be constant for relatively short periods of time. Ball's methodology will therefore have a tendency to overstate the number and understate the length of episodes. Lastly, the learning rates take as given the survey estimates of inflation expectations. It is quite possible that the available surveys are a poor proxy of longer-term expected inflation as we conceive of it here.

Figure 3

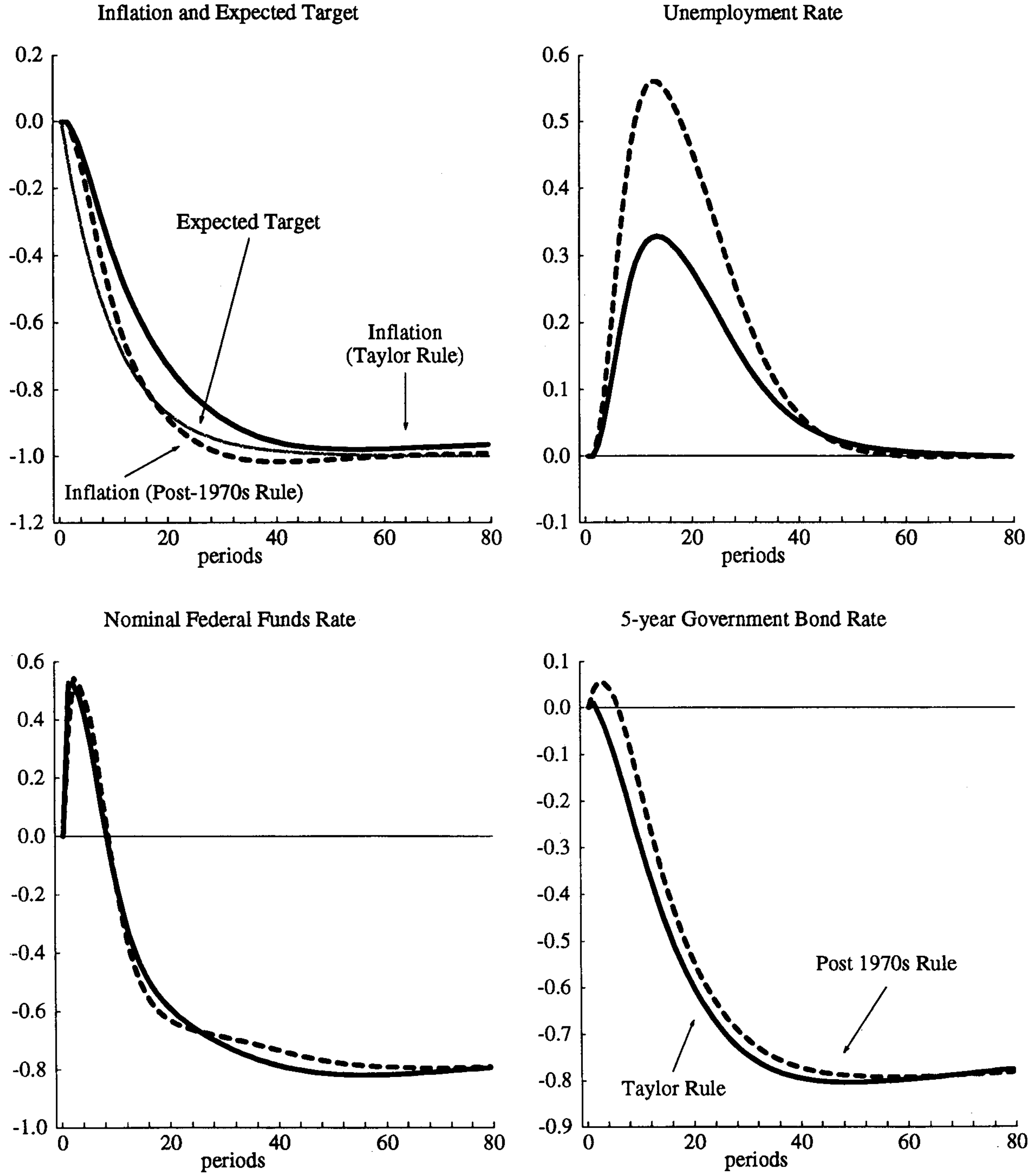
FRB/US disinflation experiments (deviations from base case)
Taylor rule versus Post-1970s rule under limited information
Learning from federal funds rate surprises at 10% rate



Now let us repeat the exercise of Table 7 for the limited-information case in Table 8. The top panel of Table 8 is identical to the bottom panel of Table 6 From Table 8 we can see that a better case can be made for a more aggressive disinflation when expectations are not model consistent, at least as they are modeled in the limited-information version of FRB/US. This is because the aggressive Post-1970s rule, given the VAR for expectations, produces larger surprises in the federal funds rate in the short run. Based on the VAR, private agents are not expecting systematically

Figure 4

FRB/US disinflation experiments (deviations from base case)
Taylor rule versus Post-1970s rule under model-consistent expectations
Learning from federal funds rate surprises at 10% rate



aggressive policy from the Fed, which is understandable given that the VAR represents the average of history; averaging tends to smooth things out.

Moreover, while private agents expect persistence in disturbances to the federal funds rate, and project declines in output on account of that expectation, the relatively mild response of the Taylor rule combined with its absence of dynamics tends to produce lower inflation than in fact materializes. The economy experiences lower output without all the disinflationary benefits. Figure 3 gives an idea of how this works; it shows a comparison of learning at 10% per quarter with the Taylor rule and the Post-1970s rule, both simulated under conditions of limited information. Notice that the aggressive action of the Post-1970s rule produces an early decline in inflation without, initially at least, a prior or coincident large increase in unemployment. This reflects the expectations effect. The decline in output, however, is more persistent so that the sacrifice ratio for the Post-1970s rule rises above that of the Taylor rule, albeit only for a short while. In this scenario, the stickiness of expectations facilitates federal funds rate surprises that aid the disinflation.

For completeness, we show in Figure 4 the 10% learning case with model-consistent expectations. Given the limited scope for perceptual errors to play a role in these scenarios, the paths of the variables tend to be quite smooth. With the obvious exception of the unemployment gap and, to a lesser extent, the inflation rates, the two rules producer broadly similar behavior. There are, however, important differences between Figure 4 and the full-credibility model-consistent simulation in Figure 2. The most obvious difference is that business cycle dynamics take considerable longer to play out when agents must learn the intentions of the Fed. In fact, the peak response of unemployment under the Taylor rule with learning (solid line, north-east panel of Figure 4) is about the same as it is under full credibility (Figure 2); unemployment simply persists substantially more when agents must learn the objectives of monetary policy. Another difference is that while the nominal federal funds rate and the government bond rate fall instantaneously in the full credibility case, they rise, initially, when agents must learn. Expected future short-term interest rates are consistently overpredicted when agents must learn which means that long bond rates are also overpredicted.

Concluding remarks

This paper has examined the issue of expectations formation and learning on the costs of disinflation, using the Federal Reserve Board's new quarterly macroeconometric model, FRB/US. Our goals in this regard were three-fold. First, we sought to demonstrate some of the properties and capabilities of the model in terms of its ability to handle meaningful policy analysis experiments that ought to be expected to engender shifts in private-sector expectations, as per Lucas (1976). In particular, we hoped to show how learning rules could be introduced into the basic model structure. Second, we wanted to examine the implications of alternative policy rules for disinflation. Our goal in this instance was demonstrative rather than exhaustive with the analysis focusing on the aggressiveness of (linear) rules, not on the fine points of specification. Finally, we wanted to examine whether the model's structure and estimation could be reconciled with measures of the historical sacrifice ratio using plausible parameters in a simple learning rule.

On our first two objectives, we would argue that the model acquitted itself well. The model is able to simulate disinflations with learning at a host of different rates and using expectations that are either model-consistent or based on limited information as represented by a small-scale VAR. Moreover, we were able to demonstrate a case for both sides in the debate on gradualism versus cold shower in disinflation.

Ironically, while proponents of cold shower policies often couch their claims in the parlance of rational expectations models, the case for cold shower disinflations in the FRB/US depiction of the US economy rests on exploiting the sluggishness of private expectations, lulled into a sense of complacency by average history policy. A measure of caution is advisable in considering this result, however, as not all cases of the Post-1970s rule under limited information better the Taylor rule, and at a horizon of 10 years, there is little to choose between the two policies at any learning rate. Nevertheless, the result is intriguing. By contrast when given what many would believe is the best possible circumstances to show gains from rapid disinflation – model consistent expectations and fast learning – our results show dominance for gradualism. Taken together, and given the myriad of other plausible learning rules that one might consider – particularly nonlinear or higher-derivative rules – one would have to conclude that the old debate of gradualism versus cold shower remains an open question.

On our final objective, our results were more mixed. If one takes for granted that the "correct" sacrifice ratio if a number of the order of 2 to 2½, measured, say, after ten years, then there are only a few combinations of expectations formation methods and rules that produce the acceptable number. An unaggressive rule like the Taylor rule will do the job, for all learning rates under model consistent expectations, and for fast learning under limited information. There are very few combinations using the Post-1970s rule that can produce a sacrifice ratio of less than 3 at 10 years or longer. There is irony in this result too since the Post-1970s rule is, as the name suggests, fitted to the Volcker disinflation, the least costly disinflation in recent history. It is possible, however, that the sacrifice ratios of Table 1 were computed early during disinflations. This possibility gives one solace since, the bottom panel of Table 8, for example, shows sacrifice ratios of 2 or less at five years, and disinflations that are between 60 and 107% complete.

References

Ball, Laurence M. (1994): "What Determines the Sacrifice Ratio?" in N. G. Mankiw (ed.) *Monetary Policy* (Chicago: University of Chicago Press), pp. 155-182.

Blackburn, Keith and Michael Christensen (1985): "Monetary Policy and Credibility: a critical review", 23, *Journal of Economic Literature*, No. 2 , June.

Blanchard, Olivier J. (1985):"Debt, Deficits and Finite Lives", 93, *Journal of Political Economy*, No. 2, pp. 223-247.

Blinder, Alan S. (1987): *Hard heads, soft hearts: Tough-minded economics for a just society* (Reading, Mass.: Addison-Wesley).

Brayton, Flint and Peter Tinsley (1995): "Polynomial Generalization of Dynamic Frictions in Structural Macro Models". *FRB staff working paper*, June.

Brayton, Flint and Peter Tinsley (eds.) (1996): "A Guide to FRB/US: A Macroeconometric Model of the United States". *Finance and Economics Discussion Series*, No. 1996-42, October.

Buiter, Willem H. and Ian Jewitt (1981): "Staggered Wage Setting with Real Wage Relativities: Variations on a Theme of Taylor", 49, *The Manchester School*, pp. 211-228.

Bullard, James B. (1991): "Learning, Rational Expectations and Policy: A Summary of Recent Research". *Federal Reserve Bank of St. Louis Review*, pp. 50-60, January/February.

Calvo, Guillermo A. (1993): "Staggered Contracts in a Utility-maximizing Framework", 12, *Journal of Monetary Economics*, pp. 383-398, September.

Chadha, Benkim, Paul R. Masson and Guy Meredith (1992): "Models of Disinflation and the Costs of Disinflation", 39, *IMF Staff Papers*, No. 2, pp. 395-431, June.

Fuhrer, Jeffrey C. (1994): "Optimal Monetary Policy and the Sacrifice Ratio", in Jeffrey C. Fuhrer (ed.) *Goals, Guidelines, and Constraints Facing Monetary Policymakers*, proceedings of a conference held at the Federal Reserve Bank of Boston, pp. 43-69, June.

Fuhrer, Jeff and George Moore (1995): "Inflation Persistence", 110, *Quarterly Journal of Economics*, No. 1, pp. 127-159, February.

Goodhart, Charles A. E. and José Viñals (1994): "Strategy and Tactics of Monetary Policy: Examples from Europe and the Antipodes", in Jeffrey C. Fuhrer (ed.) *Goals, Guidelines, and Constraints Facing Monetary Policymakers*, proceedings of a conference held at the Federal Reserve Bank of Boston, pp. 139-187, June.

Gordon, Robert J. and Stephen R. King (1982): "The Output Cost of Disinflation in Traditional and Vector Autoregressive Models". *Brookings Papers on Economic Activity*, No. 1, pp. 205-242.

Hansen, Lars P. and Thomas J. Sargent (1980): "Formulating and Estimating Dynamic Linear Rational Expectations Models", 2, *Journal of Economic Dynamics and Control*, No. 1, pp. 7-46, February.

Kozicki, Sharon, Dave Reifschneider and Peter Tinsley (1995): "The Behavior of Long-Term Interest Rates in the FRB/US Model". Unpublished manuscript, Macroeconomic and Quantitative Studies Section, Board of Governors of the Federal Reserve, December.

Lucas, Robert E. (1976): "Econometric Policy Evaluation: A Critique". *Carnegie-Rochester Conference Series on Public Policy*, No. 1, pp. 19-46, Spring.

Mankiw, N. Gregory (1991): *Macroeconomics* (New York: Worth).

Mussa, Michael (1981): "Sticky Prices and Disequilibrium Adjustment in a Rational Model on the Inflationary Process", 71, *American Economic Review*, No. 6, pp. 1020-1027, December.

Nickell, Stephen (1986): "Error Correction, Partial Adjustment and All That: An Expository Note", 47, *Oxford Bulletin of Economics and Statistics*, No. 2, pp. 199-129.

Okun, Arthur M. (1978): "Efficient Disinflationary Policies", 68, *American Economic Review*, No. 2, pp. 348-352, May.

Rotemberg, Julio (1983): "Aggregate Implications of Fixed Costs of Adjusting Prices", 73, *American Economic Review*, No. 3, pp. 343-346, June.

Sargent, Thomas J. (1993): *Bounded Rationality in Economics* (Oxford: Clarendon).

Taylor, John B. (1983): "Union Wage Settlements During a Disinflation", 73, *American Economic Review*, No. 5, pp. 981-993, December.

Taylor, John B. (1993): "Discretion versus Policy Rules in Practice", 39, *Carnegie-Rochester Conference Series on Public Policy*, No. 2, pp. 195-214, Autumn.

Taylor, John B. (1994): "The Inflation/Output Variability Tradeoff Revisited", in Jeffrey C. Fuhrer (ed.) *Goals, Guidelines, and Constraints Facing Monetary Policymakers,* proceedings of a conferences held at the Federal Reserve Bank of Boston, pp. 21-38, June.

Tinsley, Peter (1993): "Fitting Both Data and Theories: Polynomial Adjustment Cost and Error-Correction Decision Rules". FEDS *Working Paper*, No. 93-21, June.

Comments on: "Expectations, learning and the costs of disinflation Experiments using the FRB/US model" by Antúlio Bomfim, Robert Tetlow, Peter von zur Muehlen and John Williams

by Gregory Sutton

The paper by Bomfim et al. examines the macroeconomic costs of monetary policy induced disinflations within the context of the Federal Reserve Board's new quarterly macro model. In my view, this is a very interesting paper, for several reasons. First, the paper shows how the clever use of state of the art macroeconometrics can allow interesting and realistic hypotheses concerning private sector expectation formation to be incorporated into a macroeconometric model. Second, the paper provides estimates of the macroeconomic costs of disinflation under a variety of combinations of assumptions concerning the behaviour of the private and public sectors. Third, the simulations of the model reported in the paper provide insight into the debate over the choice of cold turkey versus gradualism as approaches to (monetary policy induced) disinflation.

One of the great strengths of the Federal Reserve Board's new macro model, as revealed by the myriad of disinflation experiments reported in the paper, is its flexibility. This is especially evident in the variety of assumptions about private sector expectation formation that the model is capable of entertaining. For example, the speed by which the private sector rationally updates its expectations (learns) of the central bank's long-run inflation target is a parameter of the model. The model also admits different assumptions concerning private sector expectations about the behaviour of economic variables over the short and medium-term. These expectations can be either fully rational (model consistent) or limited-information rational. In the latter case, expectations are compatible with forecasts made from a small-scale vector autoregression (VAR).

Some of the central results of the simulations of the model reported in the paper are:

1. If an announced disinflation is fully credible, so that agents immediately accept the central bank's new long-run inflation target, then a case can be made, on the basis of sacrifice ratios, for a relatively gradual approach to disinflation. The intuition behind this result is that, because of the existence of nominal rigidities, it takes time for wages and prices to adjust to the new level of long-run inflation, even if disinflation is fully credible. Therefore, a disinflation campaign that allows more time for these adjustments to be made achieves a given reduction in the rate of inflation with less foregone employment.
2. This conclusion depends on the assumption that the announced reduction in the inflation target is fully credible. If expectations are limited-information rational and the private sector is unsure as to the central bank's inflation target, then a relatively aggressive approach to disinflation may prove to be less costly in terms of foregone employment. This obtains because an aggressive approach creates larger monetary "surprises" which in turn lead private agents to revise downwards more rapidly their guess of the central bank's inflation target.

Clearly, these findings are sensible and most likely robust to minor points of model specification. Nevertheless, it would be interesting to investigate the sensitivity of the estimated sacrifice ratios to some of the maintained assumptions of the model. In this regard, an extension of the learning rule along the lines suggested by the authors would be of interest. It would also be of interest to investigate in greater depth the implications of the functional form of equation (2). This is an important element of the model, because it is intimately related to inflation dynamics and hence the estimated values of sacrifice ratios. The authors indicate that the specification of equation (2) imposes testable restrictions on the data. It would be of interest to report the results of formally testing these restrictions and of investigating the sensitivity of the simulation results to changes of the functional form of equation (2).

A structural model for the analysis of the impact of monetary policy on output and inflation

Javier Andrés, Ricardo Mestre and Javier Vallés[1]

Introduction

Monetary policy plays an important role in stabilising economic fluctuations and, especially, in controlling the rate of inflation in an economy. This paper analyses the monetary transmission mechanism in Spain by carrying out a series of simulation exercises using a macroeconomic model with a high degree of aggregation. These exercises serve to compare the behaviour of the rate of inflation and the rate of growth of GDP in the base scenario with their behaviour in various alternative monetary policy settings.

Using these simulation exercises, this study evaluates the response of inflation and the rate of growth, once all other endogenous variables (the exchange rate, the real long-term interest rate, the demand for real balances, etc.) have been appropriately adjusted. In the macroeconomic model used, no explicit distinction is made between the various components of expenditure or income. A high degree of aggregation makes it difficult to perform a detailed analysis of monetary transmission channels, but it does offer answers to some major questions that are less sensitive to alterations in the model's specifications; in particular, the response of demand to variations in the interest rate, the importance of competitiveness in monetary transmission, and the speed with which expectations are adjusted.

The simulation model's specifications take into account the new Spanish monetary policy regime adopted in the setting of European economic integration. Institutionally, that change is manifest in a stronger gearing of monetary policy towards price stability. In this context, monetary policy is understood to mean the determination of a short-term nominal interest rate consistent with the central bank's inflation target. As for the way financial markets operate, the gradual opening up of the Spanish economy and its incorporation into international markets mean that economic agents expectations play an important part in determining nominal and real variables in the economy.

The rest of the paper is structured as follows. Section 1 summarises the most salient aspects of the monetary transmission mechanism and the development of the concrete model. Section 2 presents the results of the simulations of various disinflationary strategies. The final section summarises the main conclusions.

1. Monetary policy transmission: an aggregate supply and demand model

Monetary policy affects inflation, the unemployment rate or the rate of growth of an economy by altering aggregate supply and demand. The set of channels influencing supply and demand is what is known as the monetary transmission mechanism. Before describing its main features, it is worth making two specific points. First, it is necessary to specify the environment in which monetary policy operates. Generally speaking, monetary policy has no major real impact in an

---

[1] The authors are members of the Research Department, Bank of Spain, and Javier Andrés is, in addition, affiliated to the University of Valencia.

economy in which all markets are in equilibrium and economic agents act rationally, given the immediate response of prices to excess demand. For that reason, we shall focus on the effects of monetary policy, in the first instance, on nominal variables. If, additionally, some kind of nominal rigidity exists, monetary policy will, at least temporarily, affect real magnitudes in the economy.[2]

Secondly, the relative speed with which monetary measures have an impact on prices and quantities determines the cost associated with each monetary policy strategy and hence its viability. Nevertheless, in this paper we will not study the determinants of the speed of responses to monetary impulses. The main point of reference in this regard is built into the dynamic structure of the empirical model, which conditions the results of the simulations.

The nature of the transmission channels depends on the way monetary policy is formulated. Thus, a central bank may employ strategies based on targets for a particular monetary aggregate or else it may choose to determine a path for short-term interest rates consistent with its goal of price stability or stability for any other nominal variable (the exchange rate, for instance).[3] As is generally recognised, the type of uncertainty prevailing in an economy affects optimal selection of one or other monetary strategy (see Poole (1970)). Nevertheless, the instability of the demand for monetary aggregates weakened their reputation as a reliable guide to the stance of monetary policy in many countries and relegated them to the role of important indicators only. Two reasons have been put forward in the literature to explain the short-term instability of the demand for money (Taylor (1995)). One is the development of financial markets and their deregulation. The other is the gradual opening up of economies and the consequent impact of new variables on the relationship between monetary aggregates and the expenditure decisions taken by economic agents. The following sections deal with the transmission mechanism under each of these monetary strategies.

1.1 The transmission mechanism based on a monetary aggregate

Table 1 depicts the main effects of monetary policy on the (cyclical component of) GDP and on inflation, via the various markets in an economy. Most of these effects are independent of whether monetary policy is defined in terms of a monetary aggregate or a specific interest rate, but let us assume, for the time being, that the monetary measure is understood to be the establishment of a monetary aggregate target (M_t).[4] In an economy in which the existence of some market imperfection generates insufficient demand, the effectiveness of monetary policy depends, fundamentally, on its impact on relative prices: the interest rate, the exchange rate, and real wages. Changes in the real interest rate allow alterations in the structure of demand over time, causing, in the event of a monetary contraction, a "delay" in certain consumption and investment decisions. Apart from this intertemporal substitution effect, interest rates can also have an impact on demand via their income and wealth effects. Likewise, the fact that domestically produced goods become relatively more expensive implies that companies lose market share through lower external demand. Aggregate supply (prices and wages) will also be affected by the cheapening of imported goods and by economic agents' expectations.

The most conventional monetary policy channel is the one that operates via the interest rate (*liquidity effect*). Restrictive monetary policy produces a temporary increase in the short-term nominal interest rate (usually its impact in the interbank market is immediate). In their calculations

---

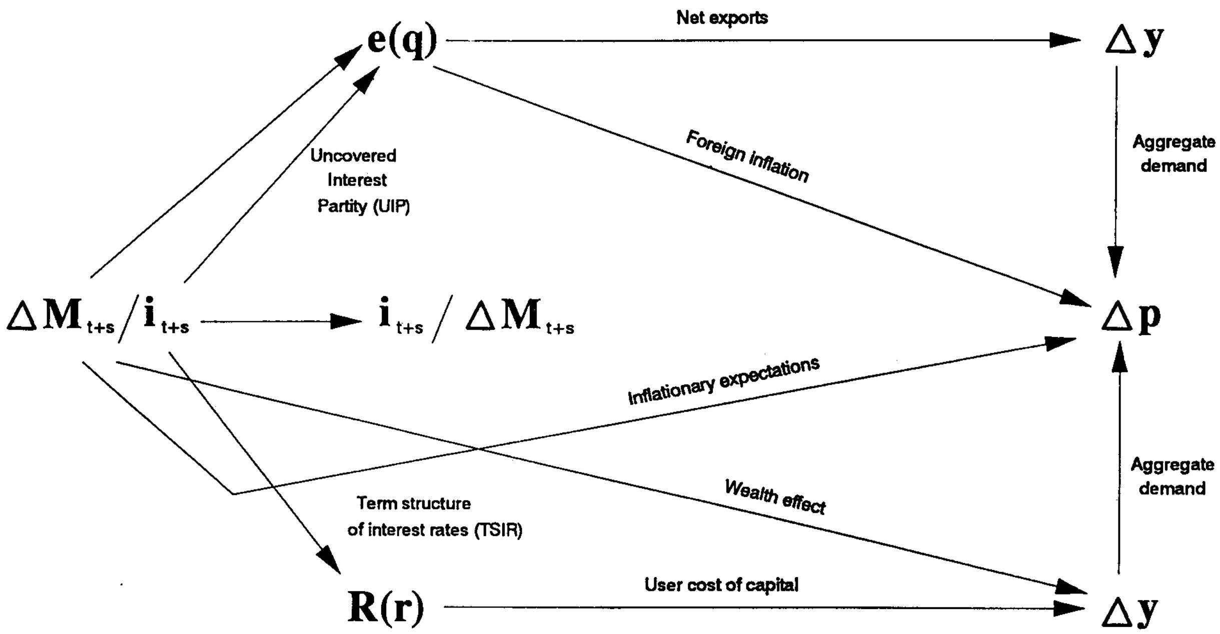
2 It is assumed that productive capacity is fixed and that it does not respond to changes in the interest rate. This being so, monetary policy has no impact on potential growth. The presence of hysteresis could generate long-term effects that are not dealt with in this paper.

3 None of these frameworks should be ascribed to a specific set of monetary policy instruments.

4 Economic agents take this value as the basis for their expectations of the future behaviour of the monetary authorities. This objective could also be defined in terms of rates of growth of the money stock.

of the user cost of capital for long-term expenditure decisions, companies and consumers use some long-term interest rate or the bank lending rate as a benchmark. The mechanism by which short-term rate signals affect the long-term rate is complex, since it is conditioned by the nature and degree of competitiveness of the financial system, the structure of short and long-term loans in corporate liabilities, and by financial markets' own perception of how temporary the change in the interest rate is likely to be. Furthermore, monetary policy can alter the relative risk premia attached to certain assets. However, monetary impulse transmission through the whole structure of assets takes time, even when the various interest rates tend to move together, and does not always proceed in a linear way.

Table 1
Money supply, output and inflation



Changes in the money supply also affect the exchange rate. How this effect operates depends notably on the degree of openness of the economy and, in particular, on the degree of international financial integration. If capital mobility is very high, the only sustainable exchange rate will be that which, risk premia aside, guarantees equality in the returns expected on domestic and similar foreign assets, when compared using the same currency. "Uncovered interest parity" is one way to represent this condition, so that the nominal interest rate differs from the expected rate depending on expectations as to the future behaviour of the interest rate differential and risk premium. Variations in a monetary aggregate affect nominal parity via three channels. First, the relative level of domestic money supply affects exchange rate projections. Secondly, the interest rate differential caused by temporary changes in the supply of real balances determines the rate of depreciation. Finally, the way monetary policy is conducted is one of the factors influencing the risk premium that investors require as an incentive to invest in domestic assets.

Generally speaking, restrictive monetary policy temporarily increases the interest rate differential, which contributes to an appreciation of the currency to an extent that creates expectations of an upcoming depreciation.[5] In addition, such a policy may lead to a reduction in the

[5] As a result of which the exchange rate would overreact and thereby attain a higher level than that pertaining after a period of depreciation.

exchange rate risk for foreign investors that helps capital inflows. Appreciation sets in motion two types of effects that help control inflation. On the one hand, the loss of competitiveness leads to a decline in net exports, and to a drop in demand pressure. On the other, the relative cheapening of imported goods reduces imported inflation and tempers wage demands.

Changes in the money supply also have an effect on aggregate supply and demand that does not operate exclusively via relative prices. Monetary policy affects the net wealth of the private sector because of its impact on the level of net assets held by that sector and on their market value via the relative price structure of financial and real assets. This *wealth effect* is the main channel in monetarist models (Metzler (1995)). There is also a direct relationship between the volume of credit to the private sector and the level of expenditure of companies and consumers. In the event of some kind of market imperfection due to asymmetric information, the marginal costs of the different forms of corporate (and consumer) financing are not equal. Specifically, a monetary policy that causes increases in the interest rate also raises the relative cost of external financing for less financially solvent firms. The *credit channel* is manifest in the restriction on the supply of bank loans associated with monetary contraction, and in the difficulty in gaining access to other external funds, both of which possibly reduce private expenditure.[6]

Finally, monetary policy may have a direct impact on inflation if the central bank is capable of influencing economic agents expectations. In the case of a restrictive monetary policy, this *expectations effect* impacts on aggregate supply, through a dampening of wage demands and of the growth in nominal costs; on the demand side, it reinforces the increase in real interest rates associated with monetary contraction. How long this increase lasts depends on how stable the drop in inflation is. As inflation falls, real balances recover with the nominal interest rate dropping in line with the new rate of inflation.

1.2 The transmission mechanism based on interest rates

The above section described the main channels through which monetary impulses are transmitted to the real economy and affect inflation. Monetary policy in Spain is more geared to a framework in which the central bank steers the current nominal interest rate (i) in accordance with its inflation target, making use of the various instruments at its disposal to make sure that the rate stays within the desired bounds.[7] In this case, i becomes the exogenous monetary policy variable and M endogenous. For a given interest rate value, the monetary transmission channels operate more or less as described in the previous section, as Table 1 shows. Nevertheless, this definition of monetary policy gives rise to at least three particularities which are worth noting.

The first refers to the type of policies that can be applied to lower inflation in an economy with higher initial levels of inflation and nominal interest rates than the countries in its immediate environment (with which it is financially integrated). Transmission of a permanent reduction in the rate of growth of the money supply operates via an increase in interest rates that may ultimately fall when inflation declines and the supply of real balances recovers. For that reason, the equivalent of a *permanent reduction* in the rate of growth of the money supply is a *temporary*, not a permanent, increase in interest rates.

This observation leads us to a second particularity in the interest rate transmission mechanism. In financially integrated economies, capital mobility tends, in the medium and long term, to make real interest rates homogeneous. Thus, in that time frame, the reduction in the inflation

---

6 Bernanke and Gertler (1995) offer empirical evidence and theoretical arguments in support of the existence of this effect. Also Hernando (1997) analyses the importance of the credit channel in monetary policy transmission in Spain.

7 See, Bank of Spain (1994):"Objetivos e instrumentalización de la política monetaria en 1995" in Boletín Económico, December.

differential is associated with a corresponding decline in the nominal interest rate differential.[8] When the central bank uses the money supply, this is achieved "automatically" if monetary contraction succeeds in lowering inflation. When the monetary authorities use the interest rate as an instrument, they can apply a policy emulating the effect of a permanent drop in the money supply by temporarily increasing the interest rate (vis-à-vis a starting point) and then reducing it later in line with the international rate. However, in theory, the possibility of an immediate and permanent reduction in the nominal interest rate, as an anti-inflationary strategy, cannot be discarded although in practice, as we argue in the following section, it turns out to be inviable.

The third main characteristic associated with a framework in which the monetary authorities control the interest rate is the difficulty of correctly determining the value of the nominal variables. When the money supply is exogenously determined, it constitutes the nominal anchor for the economy. Nominal indeterminacy with an interest rates rule appears in a wide range of models. In a closed economy with fixed prices, the price level will be determined by past history. However, when prices are fixed rationally, Sargent and Wallace (1975) demonstrated that the price level is indeterminate.

In an open economy, in which the exchange rate is determined according to rational expectations, the absence of a nominal anchor gives rise to a basic lack of determination of the exchange rate and of the other nominal variables in the economy. The main effect of this lack of determination, both theoretically and empirically, is the difficulty of correctly evaluating the cost of alternative anti-inflationary strategies. The following sections will address this point in greater detail, once the economic model has been specified.

1.3 A macroeconomic aggregate supply and demand model

We consider an open economy model, with a given level of external output and international trade.[9] The government determines fiscal policy exogenously.[10] There are four financial assets: money, short-term domestic bonds, long-term domestic bonds, and foreign assets, which means that three financial market equilibrium conditions are required to determine the relative prices of those assets. The short-term interest rate is established by the monetary authorities. The model can be expressed as a function of the following set of equations showing the behaviour of aggregate demand[11] and supply:

$$y_t - y_t^p = \alpha_{10} + \alpha_{11} r_t + \alpha_{12} q_t + \alpha_{13} g_t + \alpha_{14} a_t + \alpha_{15} y_t^* + \varepsilon_{yt} \qquad \text{IS (1)}$$

$$m_t - p_t = \alpha_{20} + \alpha_{21} R_t + \alpha_{22} i_t + \alpha_{23} y_t + \varepsilon_{mt} \qquad \text{LM (2)}$$

$$r_t \equiv R_t - E\left[\Delta p_{t+1/t}\right] \qquad \text{RIR (3)}$$

$$R_t = (1-b) i_t + bE\left[R_{t+1/t}\right] + \varepsilon_{it} \qquad \text{TSIR (4)}$$

$$i_t = i_t^* + E\left(e_{t+1/t}\right) - e_t + \varepsilon_{et} \qquad \text{UIP (5)}$$

---

8 The real interest rate is determined at the international level. In a closed economy, it also depends on preferences and technology, but it is possible to alter it in the short term.

9 See Dornbusch (1976), Duguay (1994), Buiter and Miller (1981) and Taylor (1995).

10 To simplify the model, one fiscal policy indicator is defined and the public sector budget restriction is not explicitly included. The interaction between monetary policy and fiscal policy is discussed, for example, in Canzoneri and Diba (1996) or Marín and Peñalosa (1997).

11 The variables in equation (1) should be understood as deviations from their long-term level or trend.

$$\Delta P_t^d = \alpha_{60} + \alpha_{61} E\left[\Delta p_t^d\right] + a_{62} E\left[\Delta e_t + \Delta p_t^*\right] + \alpha_{63}\delta\left(y_t - y_t^p\right) + \varepsilon_{pt} \qquad \text{AS (6)}$$

$$\Delta p_t = \alpha_{70} + \alpha_{71}\mu_t + \alpha_{72}\Delta p_t^d + \alpha_{73}\left(\Delta p_t^* + \Delta e_t\right) + \varepsilon_{pt} \qquad \text{CPI (7)}$$

The endogenous variables are: cyclical output[12] ($y_t - y_t^p$), the change in domestic prices measured by the GDP deflator (Δp_t^d), the money stock in real terms ($m_t - p_t$), the long-term nominal interest rate (R_t), the real interest rate (r_t), the exchange rate (e_t), and the inflation rate (Δp_t) as measured by changes in the CPI. The exogenous variables are: potential output (y_t^p), the fiscal policy indicator (g_t), world income (y_t^*), the short-term interest rate (i_t), external inflation (Δp_t^*) and the share of imports in income (μ_t). The exchange rate is defined as the value of the foreign currency in local currency terms, and q_t is the real exchange rate. The definition of wealth (a_t) only includes real balances for lack of other quarterly asset figures. All the variables are expressed as logarithms, except the interest rates. We assume that the expectations in the financial equations ((3), (4) and (5)) are fully anticipated by economic agents. Conversely, the supply-side rigidities (equation (6)) imply that expectations are formed by past values of the explanatory variables. Before describing the estimations of the equations in the model, we shall briefly explain the specifications chosen for each of them.

Equation (1) represents aggregate demand as a function of the long-term real interest rate, the real exchange rate, a fiscal indicator, the level of wealth, and world income. Thus, changes in relative prices operate through three channels in aggregate demand: the substitution effect between consumption and saving and between domestic and foreign goods; the negative income effect for economic agents who are net creditors; and, finally, a wealth effect which increases the value of residents' assets in the event of declines in the interest rate or depreciations of the local currency.

The econometric evidence for Spain reveals a negative effect of real interest rates on the components of final demand for goods and services. That is the finding of estimations made using annual investment functions (Andrés et al. (1992)) and consumption functions (Estrada (1993)). Similar evidence is obtained using quarterly data on the demand for durable and nondurable consumer goods and on the demand for domestic housing and private sector productive investment (see Estrada et al. (1997)). The results of these calculations reveal a moderate impact of the real interest rate (compared to the effect of other variables) appearing after a certain time lag. This significant impact of the real interest rate on consumption and investment is not to be found in the same degree of intensity on the more aggregate magnitudes: output and inflation. The available evidence suggests that short-term interest rates are much less of a guide to the level of economic activity and price behaviour than changes in the money stock (Alvarez et al. (1995)). Our model reflects the direct impact of the money stock by including the stock of liquid assets held by the public (ALP) in the determinants of aggregate demand.[13]

The real exchange rate is an important variable in analysing the monetary transmission mechanism. The gradual opening up of the Spanish economy has increased the possibilities of substituting external demand for domestic demand depending on how our prices behave vis-à-vis

---

12 For which the rate of growth of GDP would be another approximation. In this case, however, the model is not neutral in the traditional sense because monetary policy affects the level of GDP in the long term as well as its composition. Nevertheless, after the adjustments in the nominal variables, the economy recovers the rate of growth considered to be that pertaining to a steady state.

13 However, this inclusion can be interpreted in two different, albeit not mutually exclusive, ways. We might be dealing with a credit channel in monetary transmission or else with a wealth effect reflecting the influence of changes in monetary policy on the relative prices of the set of assets.

those abroad.[14] Moreover, in an economy that is fully integrated with the outside world, the possibilities of influencing aggregate demand in the short term are not just limited to the real (after-tax) interest rate: they include the exchange rate as well. In the long term, nevertheless, the real interest rate will be determined by that prevailing abroad and the real exchange rate will be determined by structural parameters in the economy.

The other two variables determining aggregate demand are external output and the government's fiscal position. The variations in external output may be reflected in total demand via declines or increases in exports. At the same time, it is to be expected that the public deficit has a direct impact on aggregate demand, quite apart from indirect effects via the interest rate[15] and competitiveness.

As for the other equations, the demand for real balances (equation (2)) depends on the level of income and the two types of nominal interest rates: short and long-term. Equation (3) defines the real interest rate. Equations (4) and (5) complete the financial structure of the model and relate the interest rate controlled by the monetary authorities to those interest rates variables and exchange rates that influence economic agents' expenditure decisions.

Variations in the intervention interest rate are initially transmitted to the money market and from there to the equilibrium rate in the lending and deposit markets. According to Ayuso et al. (1994), the interbank rate in Spain responds fully and relatively swiftly to shifts in the intervention rate. That finding justifies using a short-term interbank rate as an exogenous rate.[16] The transmission of monetary policy signals to the rates that determine spending by economic agents will depend on the structure of the financial sector. Among other determining factors, of particular importance are the liquidity and degree of maturity of the different assets and the degree of competition among financial institutions. The fact that much of the spending by economic agents in Spain depends on external funding and that the composition of bank assets is sensitive to monetary policy could lead one to consider the lending rate of financial institutions as a relevant interest rate in expenditure equations. Nevertheless, in this paper we have opted to take the long-term debt rate as an indicator of the cost of capital in consumption and investment decisions, owing to the similar influence that both variables exert in the demand equation. The term structure of interest rates (equation (4)) is represented in the model as an arbitrage condition between short and long-term nominal interest rates: the long-term interest rate being the present value of expected short-term interest rates.

Our model uses the uncovered parity of interest rates (equation (5)) to reflect the way monetary policy affects the exchange rate. This implies that perfect capital mobility is the best way to represent the degree of financial integration of our economy with the outside world. The error term in this relationship allows for possible temporary shifts in the exchange rate due to the emergence of risk premia or speculative runs in exchange markets that are unrelated to macroeconomic variables.

We shall assume that the financial sector makes rational predictions of future changes in monetary policy. The developments expected in short-term interest rates determine both the long-term interest rate – through the term structure of interest rates – and the behaviour of the exchange rate through uncovered interest rate parity. The inclusion of this arbitrage condition has some additional advantages. First, it allows us to attribute a predominant role in monetary transmission to economic agents' expectations regarding future monetary policy. The second advantage is that this

---

14 This appears to be an important variable in the Spanish economy, since the effective exchange rate is one of the main determinants of net exports (see, for instance, Buisán and Gordo (1993) and Bajo and Montero (1995)).

15 In Spain, the evidence in this regard is not conclusive. Thus, whereas Ballabriga and Sebastián (1993) do not find a causal relationship between interest rates and the deficit, Raymond and Palet (1990) do report a direct effect.

16 For the purposes of this paper, we prefer this to the alternative of endogenising the interest rate as a function of the reaction of the monetary authorities (see, for instance, the estimations by Escrivá and Santos (1991)).

obviates the need for a more complex model, which would have to incorporate other assets and a study of the productive structure of the financial sector.

Aggregate supply in the economy is represented by equations (6) and (7). Equation (6) is a Phillips curve which relates domestic inflation to inflationary expectations and demand pressure. Inflationary expectations are represented in the model as a function of past price variations, for both domestic and foreign goods. Inclusion in the equation of a variable reflecting the business cycle situation of the economy makes it possible to measure the response of prices to excess supply or demand in goods markets. Equation (7) is a dynamic version of the consumer price index definition. In the long-term solution of (6) and (7), nominal homogeneity is assumed.[17] The current version of the model does not include possible supply shifts, such as variations in raw materials prices, changes in indirect taxation, or deviations in the relationship between wages and productivity. For the sake of simplicity, we have also eschewed incorporating a set of price and wage equations, even though this means sacrificing data regarding determinants of the degree of nominal rigidity influencing monetary transmission.

2. Monetary policy simulations

2.1 The estimated model

The estimated long-term version of the model is presented in Table 2, and the short-term estimations are to be found in the Appendix A. The explanatory variables were selected on the basis of lags in those same variables and of contemporary and past values of the other endogenous and exogenous variables. The maximum number of lags considered was eight, the idea being to register correlations within and between annual periods. The selection criteria were the significance of each variable, on its own and as part of the set, and its stability. Each equation was estimated using instrumental variables for the sample period 1970Q1-1994Q4. Past values of all the endogenous and exogenous variables of the model were taken as instruments.[18]

When estimated, equation (1), corresponding to IS, shows a persistent rate of growth of output over four periods. The external output cycle (y^*) affects Spanish output after a two-quarter lag. While the real rate of exchange (q) has an impact on aggregate output in the same period, an increase in the long-term real interest rate (r) only contracts demand after a seven-quarter lag. The wealth effect is reflected here in the rate of growth of real monetary balances (m/p). This wealth effect is represented in the model by excluding assets abroad and making a separate estimation of demand for ALP.

The fiscal impulse variable in the IS equation is measured according to the second differences in public consumption (g).[19] The econometric specifications chosen limited the dynamics of variable g in such a way as to restrict its long-term elasticity in the model to zero.[20] There are various reasons for restricting the fiscal variable to have only a transitory impact. First, the sample

---

17 So that an increase in domestic inflation and in imported inflation gives rise to an equivalent increase in the CPI inflation rate.

18 These estimations would have been more efficient if the residual correlation between equations had been taken into account (estimating 3SLS, for instance). We will leave that task for future research.

19 This is the only variable available in Quarterly National Accounts.

20 This differs from the elasticity estimated for the sample period (0.6), which coincided with a period in which the growth of government consumption far exceeded that of GDP.

period was characterised by a steady increase in the ratio of government consumption to GDP that is unlikely to be repeated in future. Secondly, we know that shifts in government consumption are not the only relevant variable for measuring all the effects of the government's budgetary restrictions. Finally, the fiscal policy exercises carried out in the unrestricted model generate a very high and implausible value[21] for the fiscal multiplier.

Table 2
The estimated model: long-run solution

$\Delta y_t = 1.89\Delta^2 g_t - 0.08\Delta r_t + 0.11\Delta q_t + 0.33\Delta(m_t - p_t) + 0.47\Delta y_t^* + \varepsilon_{yt}$	IS (1)
(4.4) (1.2) (3.3) (4.8) (2.5)	
$m_t - p_t = -5.5 - 0.03\Delta R_t - 0.15\Delta i_t + 1.25 y_t + \varepsilon_{mt}$	LM (2)
(0.8) (1.7) (23.5)	
$r_t \equiv R_t - E[\Delta p_{t+1/t}]$	RIR (3)
$R_t \equiv (1 - 0.91) i_t + 0.91 E[R_{t+1/t}] + \varepsilon_{it}$	TSIR (4)
(*) (13.4)	
$i_t = i_t^* + E(e_{t+1/t}) - e_t + \varepsilon_{et}$	UIP (5)
$\Delta p_t^d = 0.98 E[\Delta p_t^d] + 0.02 E[\Delta e_t + \Delta p_t^*] + 0.25\Delta y_t + \varepsilon_{pt}$	AS (6)
(1.5) (*) (3.7)	
$\Delta p_t = -0.57\Delta^2 p_t^d + 0.14(\Delta^2 p_t^* + \Delta^2 e_t) + \Delta P_t^d + \varepsilon_{pt}$	CPI (7)
(3.4) (3.7) (*)	
(*): restricted	

Definition of endogenous variables:
Δy_t: rate of growth of GDP; Δp_t^d: GDP deflator inflation; Δp_t: CPI inflation; e_t: nominal exchange rate (q_t real exchange rate); R_t: nominal long-term interest rate; r_t: real interest rate; $m_t - p_t$: supply of real balances.

Definition of exogenous variables:
g_t: government consumption; y_t^*, world income; i_t: short-term interest rate; i_t^*: external short-term interest rate; Δp_t^*: external inflation.

In the demand for real balances (equation (2)), (m-p) corresponds to a broad monetary aggregate, ALP2, which includes commercial paper as well as short-term public debt. This is the aggregate used as a relevant indicator by the Banco de España for monetary planning purposes. A long-term relationship is estimated between income and real balances, which implies an elasticity somewhat greater than one. The short-term variables are, apart from a lag in the variable itself, lags in the long-term nominal interest rate (R) -which measures the return on assets other than money – and accelerations in the inflation rate ($\Delta^2 p$). The three-month interbank rate (i) figures weakly in this relationship, also with a minus sign. This may measure the effect on demand of a change in the money supply that is not reflected in longer-term rates. For reasons of simplicity, such demand for money is not assumed to reflect the interaction between the return on assets included in ALPs and

21 Compared to the values generated by other models, such as the NIGEM, for the Spanish economy and economies similar to ours.

the return on alternative assets that in earlier estimations proved to be important for the Spanish economy (Cabrero et al. (1992) and Vega (1997)).

A linear approximation has been estimated for the arbitrage condition between the interbank interest rate and the long-term debt rate (equation (4)). The parameter estimated compares the return on two assets with different maturities, assuming risk neutrality and rational expectations by economic agents.[22] The existence of a relationship of cointegration between short and long-term interest rates was tested and accepted, and parameter *b* of equation (4) was subsequently estimated, with a value very similar to the steady-state value for the long-term interest rate obtained for other countries (Taylor (1993)). Equation (5), which assumes interest rate parity, incorporates a risk premium.

Equation (6) on domestic prices (Δp^d) proxies inflationary expectations by lagged values of foreign and domestic prices, and demand pressure by the lagged coefficient of GDP growth. Price accelerations appear as additional regressors. This calculation yields a figure for the impact of external prices of 0.03, for a sample period in which the Spanish economy opened up considerably, so that a much higher figure is likely in the future with a more integrated economy. For the simulation, we took a higher figure (0.13), which, in our opinion, is closer to the current weight of imports in the national economy. This greater responsiveness of economic agents to expected changes in external prices accelerates the convergence between domestic and foreign variables.

The last equation estimated in Table 2 deals with the consumer price index (Δp), for which cointegration with domestic prices was found. The equation was therefore estimated taking ($\Delta p - \Delta p^d$) as the explanatory variable. External prices have only a transitory effect on the consumer price index and their permanent impact on inflation will be via expectations of domestic prices. The long-term restrictions contained in this equation (coefficient of 1 for domestic prices and 0 for external prices) are accepted statistically.

2.2 The determination of the nominal magnitudes and the monetary policy strategy

The financial structure of the model described in the previous section allows us to analyse the behaviour of the central bank in its goal of controlling inflation by establishing an interest rate path. Despite its simplicity, this framework underscores certain basic issues that the central bank faces in designing monetary policy and which affect the use of the model for simulation purposes.

The long-term solution of equations (1) - (7) can be represented as:

$$i = i^* + \Delta e \qquad (8)$$

$$\Delta p = \Delta p^* + \Delta e \qquad (9)$$

this last condition reflecting that the economy cannot systematically gain or lose competitiveness.[23] Equations (8) and (9) determine the level of the long-term real interest rate, which should be at the international level. Hence, reducing domestic inflation to the international level requires setting the domestic interest rate (i) at the international level (i^*).

$$i - \Delta p = i^* + \Delta p^* \qquad (10)$$

In light of this long-term condition, one query concerns the path that the domestic interest rate must follow from a starting-point in which the economy has a positive inflation

22 See the article by Ayuso and Núñez (1997) to obtain a term structure with zero coupon assets.

23 In addition, the output gap has to close in the long run.

differential vis-à-vis external inflation. As pointed out earlier, an immediate and permanent reduction of the nominal interest rate cannot be dismissed out of hand. A central bank move in this direction gives rise to an immediate expansion, which in turn produces a generalised loss of competitiveness via an initial increase in inflation. If the central bank were able to maintain this policy indefinitely, sooner or later the deterioration in the trade balance would end up producing a recession and, along with it, a decline in inflation. Now, in practice, a temporary increase in inflation may erode the credibility of the ultimate objective, which makes this strategy completely inviable. This suggests the need for a more suitable, in the end equivalent, monetary strategy (that is: $\Delta p = \Delta p^*$), which involves a temporary increase in the domestic interest rate – thereby encouraging a contraction – and then a gradual reduction to its long-term level as the inflation goal is attained. Unlike the previous case, the goal of controlling inflation is attained more rapidly, although the costs in terms of output and employment are brought forward in time.

The existence of two quite opposite interest rate paths that lead, nevertheless, to the same long-term objective, indicates that there are indeed numerous intermediate alternatives, the relative advantages of which have to be assessed. In order to make an assessment of these or other economic policy strategies, it is necessary to know how the nominal exchange rate reacts to shifts in the economic environment. The greater the sensitivity of the nominal exchange rate to interest rate movements, the more intensely inflation will respond in the short term to changes in the nominal interest rate. What does the model represented in equations (1) - (7) tell us about the behaviour of the exchange rate when confronted with a change in the pre-announced interest rate path? Unfortunately, not enough. Fulfilment of the uncovered interest rate parity condition indicates that the current exchange rate value differs from the expected value by an amount equal to the interest rates differential, but that tells us nothing about the expected value itself. This value is determined by, among other factors, the level of the domestic money supply in relation to the supply of foreign currency. When the central bank fixes the interest rate, the money supply is endogenous and how it moves depends on the behaviour of the price level, which in turn depends on the behaviour of the exchange rate. In this way, the level of the nominal variables remains undetermined and it is not possible to predict the response of the exchange rate to a specific monetary policy measure (see Appendix B). Thus, while it is possible to calculate the exact level that real variables will reach, the existence of multiple solutions for the nominal variables implies that the transition path between two long-term equilibria is not unique.

Nominal indeterminacy is a feature of a wide range of models. A solution to this problem involves incorporating another equation into the model which, in one form or another, fixes the value of some nominal value. The generic form of such an equation is that of a feedback rule, according to which the monetary authorities react by varying the interest rate when some variable (X) deviates by more than a desired margin from its target value.[24] This way of constructing the model nevertheless poses some problems, particularly in respect of the choice of target variable.[25] If X stands for the *price level or parity of the peseta*, the feedback rule could be understood as an equation fixing a target of the monetary authority in terms of these variables (in either direction).

[24] This type of rule takes up McCallum's proposed solution to the non-uniqueness problem. Application of monetary policy by fixing an intermediate objective in quantitative terms would indeed be a specific instance of such rules, in which X is the money supply.

[25] At the same time, inclusion of a feedback rule appears to be logical in a model representing, from the outside, the behaviour of the monetary authorities, but not as an element to be taken into account in monetary policy decisions. Moreover, choice of an optimal rule implies first tackling the ever-complex question of the definition of a target function of the monetary authorities.

However, this would be an inappropriate representation of monetary policy in Spain.[26] For the simulation exercises we discuss below, a further condition has been added to the model. Specifically, it is assumed that an immediate reduction in interest rates leads to a depreciation of the peseta. Conversely, the peseta appreciates when the central bank raises interest rates. This assumption does not, however, solve the issue of how large is the magnitude of the exchange rate response. We will return to this point in our description of the simulation exercises.

2.3 Simulations

This section will assess the relative importance of the channels through which monetary policy operates by simulating the impact associated with alternative interest rate (and exchange rate) paths. These exercises have not been carried out in order to lay down how the Banco de España's monetary policy should be implemented; our goal is more modest and restricted to analysing within a given framework the importance of certain monetary policy transmission mechanisms. Consequently, no attempt has been made to establish optimum inflation targets or plausible ranges of fluctuation around them.[27] All the simulations are compared with a hypothetical scenario, independent of the model itself, in which it is assumed that domestic and foreign economic activity grows by 2% a year, domestic prices by 5% per annum, and external prices by 3% annually. The simulations reflect the transition between an equilibrium with 5% inflation to another with 3% inflation. The object of the exercise is to establish the procedures and costs of eliminating the inflation differential vis-à-vis the rest of the world. The linearity of the model, at least with regard to the variables expressed as logarithms, makes it possible to isolate these responses from the baseline, which thus becomes of secondary importance. The simulations were carried out by solving forward the financial variables, which means that the values for those variables anticipated by economic agents coincide with what was actually simulated. All the equations employ residuals, in such a way that each equation reproduces, for each period, the data given in the scenario.

Given the structure of the model, the 2-point reduction in domestic inflation requires an equivalent, long-term, reduction in the domestic interest rate. First, we shall simulate the effects of an *immediate and permanent reduction of interest rates of 2-points* (simulation 1). As discussed in the previous section, the instantaneous response of the exchange rate has to be imposed on the basis of criteria independent of the model. A reasonable hypothesis is that the 2-point lowering of interest rates goes together with an initial nominal exchange rate depreciation of 5%.[28] Charts 1 and 2 show the behaviour of inflation and GDP growth after application of this policy. The most notable feature of this simulation emerges in the inflation rate profile for the time period under consideration. Although this monetary policy measure ultimately leads to the targeted level of inflation, the reduction in interest rates and the nominal depreciation produce a temporary increase in inflation which may endanger the disinflation process itself. Despite the modest exchange rate depreciation, its short-term impact on inflation is considerable: for over a year, inflation remains higher than 6% and only starts declining after three years.[29] This effect is due to the increase in external prices, which swiftly and strongly raises inflationary expectations and affects demand pressure, albeit not to the same extent.

---

26 Nominal stability targets are generally expressed in terms of the inflation rate, which is compatible with very different price level values. Establishing an *inflation rate* target does not solve the indeterminacy problem (Dhar et al. (1994)).

27 See, for example, Haldane et al. (1995).

28 All these figures exclude the risk premium (which is exogenous) and are, moreover, approximate, given that the original movements are expressed as logarithms.

29 While generating a small expansion of GDP for two years (Chart 2).

Chart 1
Alternative interest rate strategies
Inflation (CPI) rate, annualised

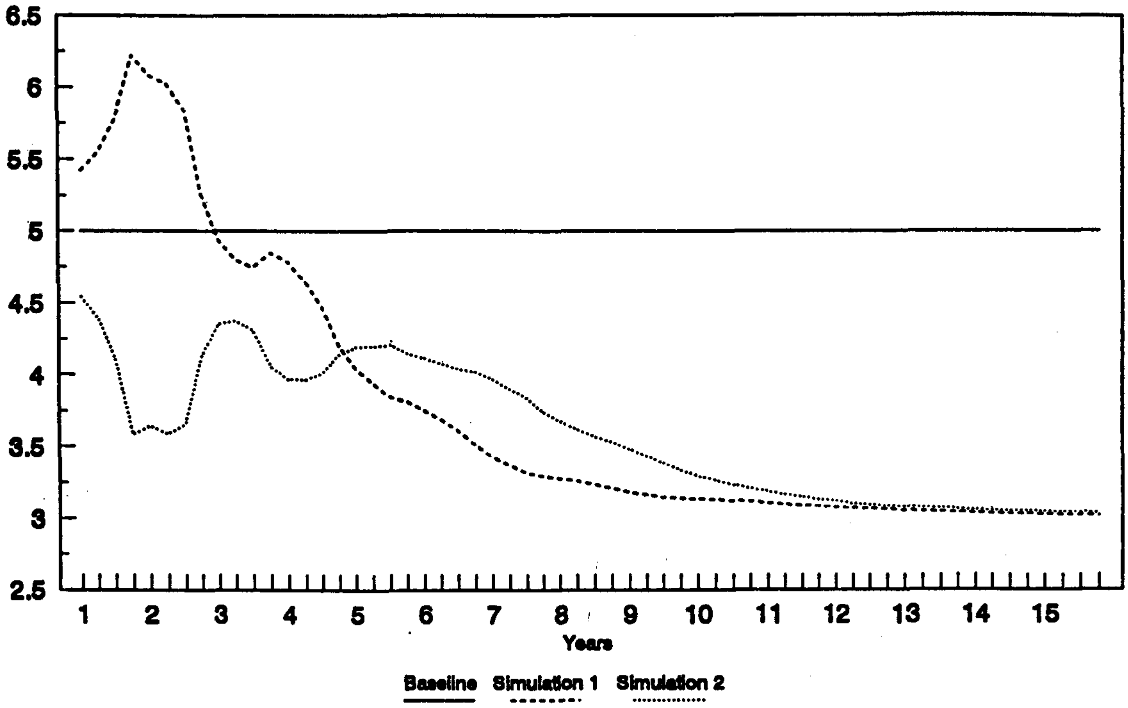
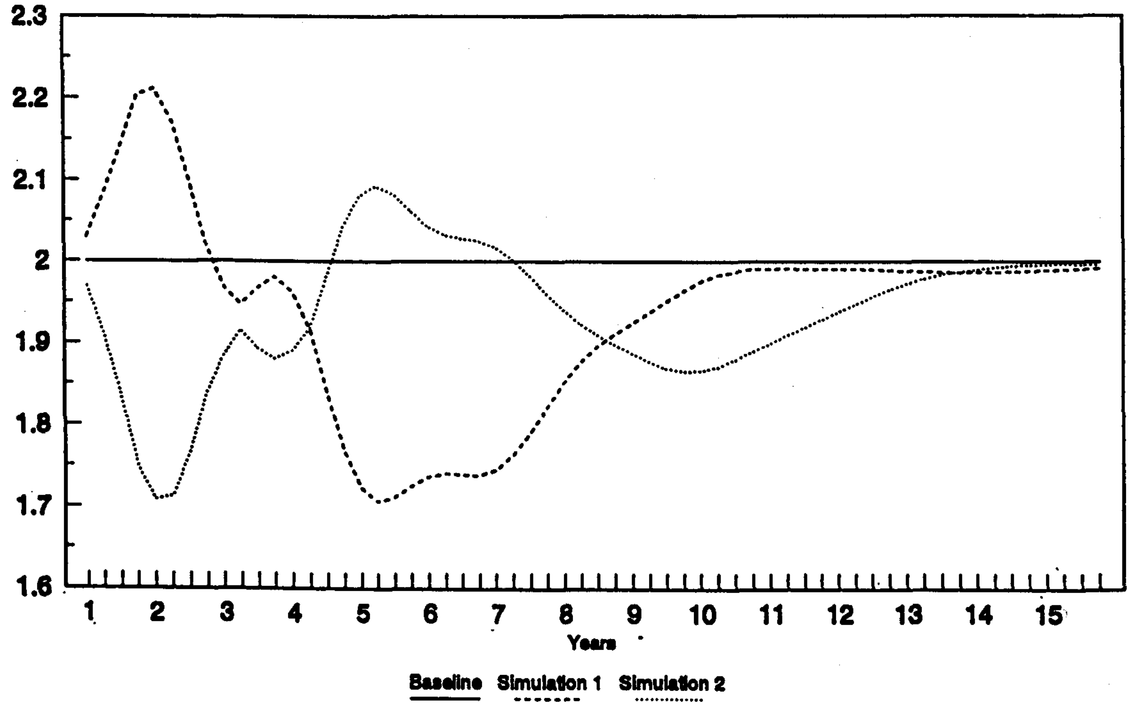


Chart 2
Alternative interest rate strategies
GDP growth, annualised



Simulation 1: Immediate reduction of interest rate with a 5% depreciation.
Simulation 2: Temporary increase, with a 5% appreciation.

Chart 3

Simulation of monetary policy: behaviour of the interest rate and exchange rate

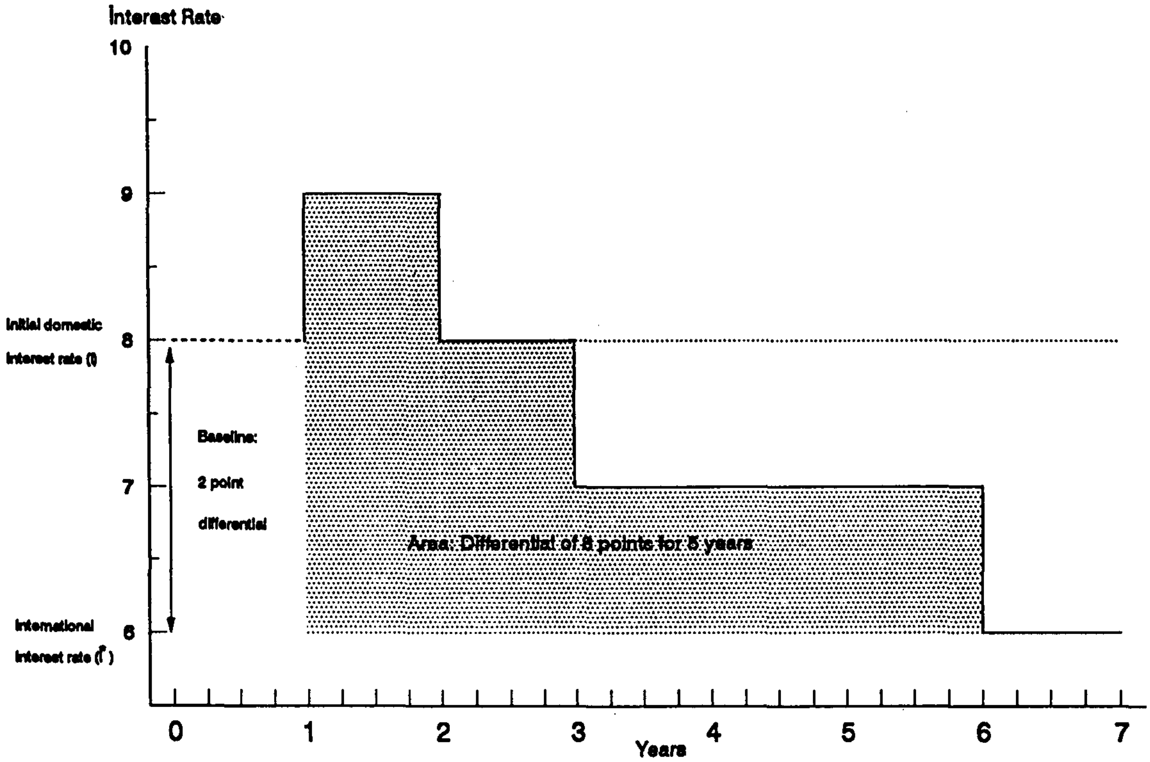
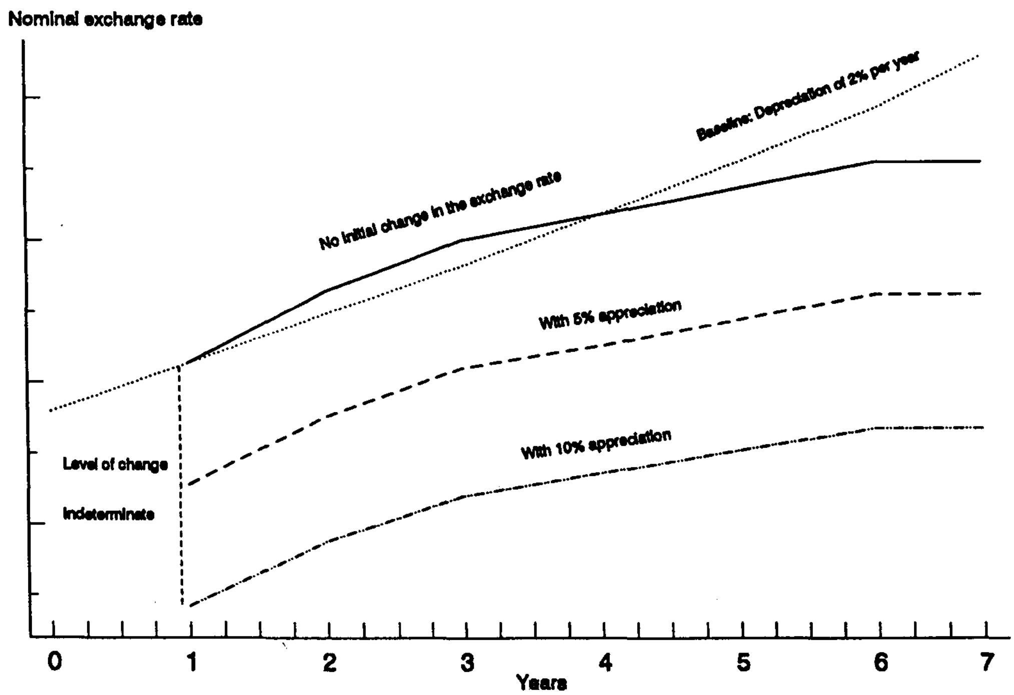


Chart 4

Exchange rate consistent with the interest rate differential



The initial depreciation of the currency is, therefore, inflationary. How does a decline in inflation take place later? The elimination of the interest rates differential over time leads the nominal exchange rate to stabilise around the new value obtained after the 5% drop on impact. Thereafter, on the supply side, imported inflation drops immediately from the 5% envisaged in the base scenario[30] to 3% (the pace at which external prices increase), thereby helping to dampen inflationary expectations and, thus, domestic inflation. Moreover, with domestic prices still rising faster than those of our competitors, our products become more expensive, which reduces external demand and sparks off a temporary recession, which also moderates inflation. Reduction of the nominal interest rate works in the opposite direction, but this effect is not sufficient to offset the loss in competitiveness. Later on, the reduction in inflationary expectations allows the real interest rate to recover and this, too, helps moderate inflation.[31]

A reduction in the nominal interest rate ensures, in principle, a reduction of inflation in the long run, but if, as a result of the initial currency depreciation, the process of transition requires a higher inflation than that prevailing at the outset, a *legitimate doubt* arises as to the *credibility* of this policy. If economic agents are sceptical about achieving the objective of ultimately reducing inflation and consider that this policy is not sustainable, it may take time for inflation to fall or it may not do so at all, and this in turn would prevent the planned reduction in nominal interest rates. The policy has to be credible. One way to minimise these risks is to effect *a temporary increase in interest rates, before reducing them permanently* to the international level. Thanks to this temporary increase, there is less risk of generating higher inflation in the short and medium term, which strengthens the credibility of the anti-inflationary policy itself. Let us observe how this policy works in the context of the model.

Take a monetary policy measure which consists of raising by 1 point the intervention interest rate for one year and then returning the following year to the level of the base scenario; the interest rate is then reduced by 1 point for another three years, ending up with a permanent reduction of 2 points. Charts 3 and 4, respectively, show the behaviour of the nominal interest rate and the nominal exchange rate, making three alternative assumptions regarding the appreciation in impact.

In accordance with uncovered interest parity, this exercise unequivocally determines the behaviour of the nominal exchange rate. This is consistent with an 8-point depreciation during the years of the transition (equal to the sum of the interest rates differentials during the first five years of the simulation), stabilising when the domestic interest rate is at the international level. However, given nominal indeterminacy, it is not known at what level the exchange rate is established. To solve this indeterminacy, an instantaneous *appreciation of 5%* is assumed to accompany the increase in the interest rate (simulation 2).[32] The findings of simulation 2 are also shown in Charts 1 and 2 to facilitate comparison. Charts 5 and 6 show the behaviour of the real interest rate and the real exchange rate in both simulations. Inflation declines in the first year by a little more than 1 point, stabilising at around 4% for the following 5 years, and beginning thereafter a gradual decline to its new long-term level of 3%. The first year, GDP grows by ¼% less than in the base scenario, gradually recovering thereafter to reach its long-term growth level in 4 years, followed by a small cycle.

---

30 Resulting from 3% external price inflation, plus 2% permanent depreciation of the domestic currency.

31 The behaviour of the real interest rate and the real exchange rate are reflected in Charts 5 and 6, which we comment on below.

32 The history of the variable itself can give us an idea of the jump: only very occasionally has the currency appreciated or depreciated by more than 5% in nominal terms (on 6 occasions in our sample period). In fact, this jump is almost twice the standard deviation in the nominal exchange rate. Almost the same is true of the real exchange rate which has only occasionally registered variations of that magnitude within a single quarter.

Chart 5
Real interest rate

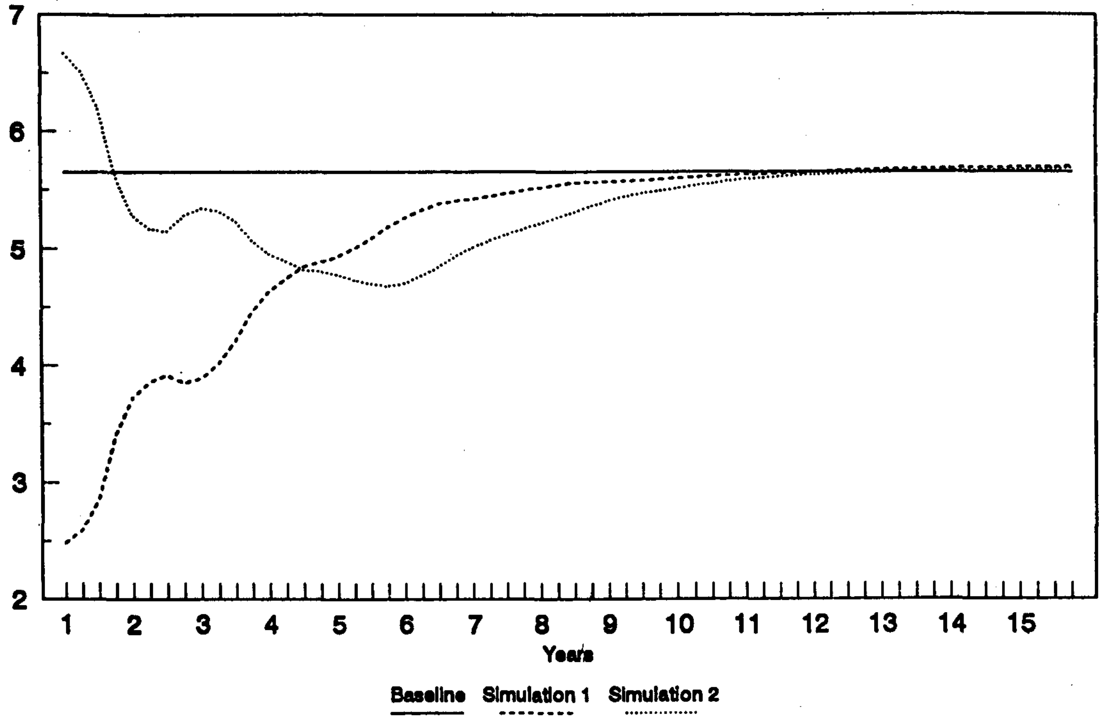
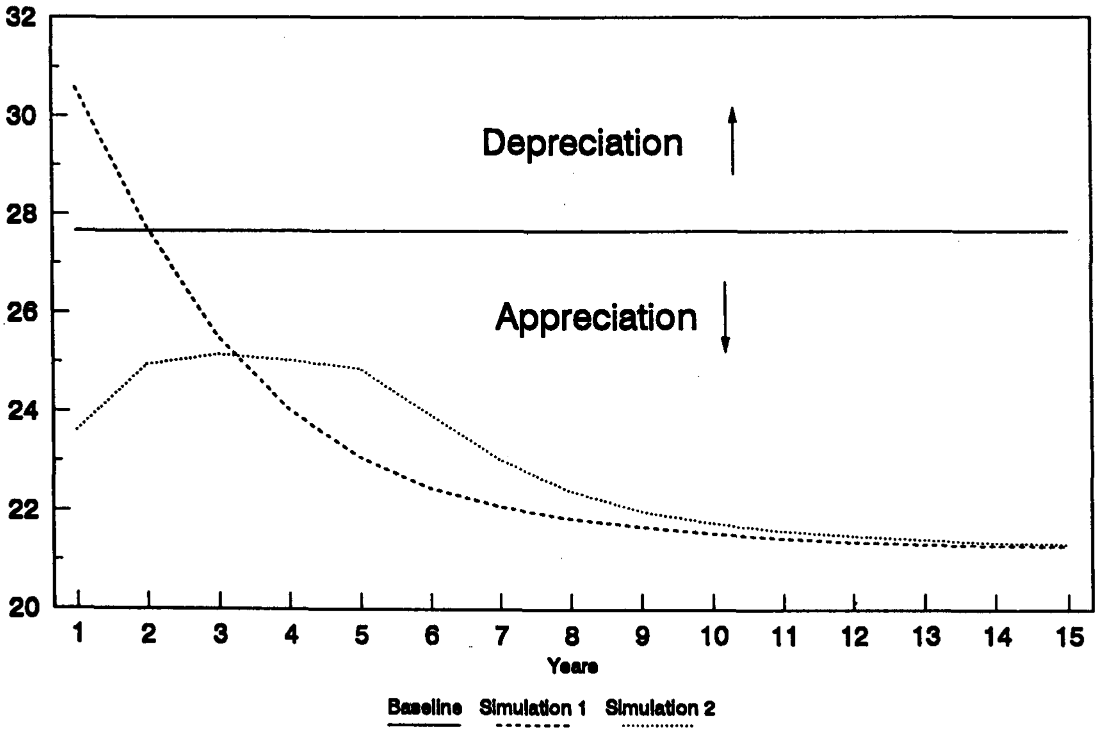


Chart 6
Real exchange rate



Simulation 1: Immediate drop and 5% depreciation.
Simulation 2: Temporary increase and 5% appreciation.

Chart 7

Temporary increase in the interest rate, dropping subsequently to the international level
Inflation (CPI) rate, annualised

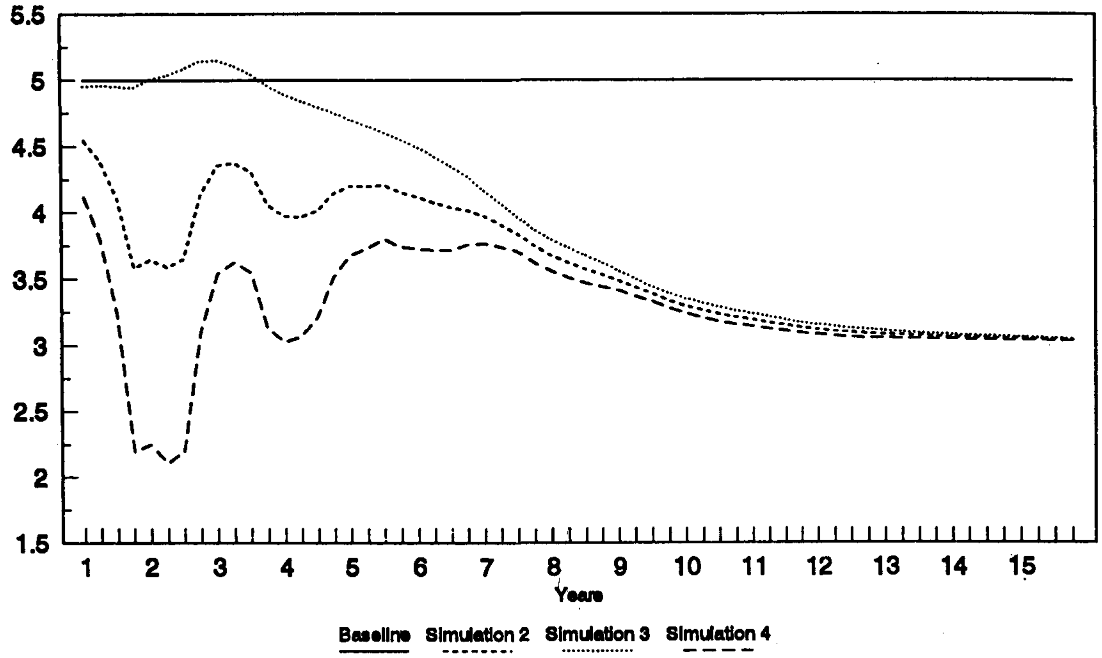
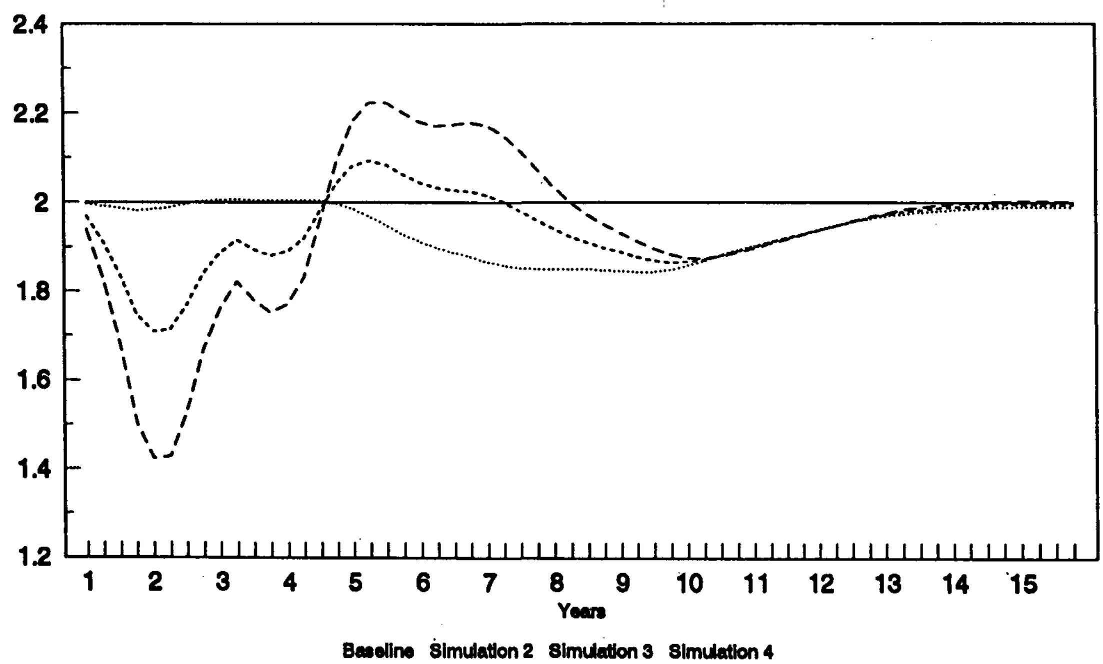


Chart 8

Temporary increase in the interest rate, dropping subsequently to the international level
Growth of GDP



Simulation 2: Temporary increase with initial appreciation of 5%.
Simulation 3: Temporary increase with no initial change in the exchange rate.
Simulation 4: Temporary increase with initial appreciation of 10%.

The initial drop in inflation is mainly due to the (expected) cheapening of imported goods: a supply effect that dampens inflationary expectations. The tiny recession thereby induced serves to keep inflation below its initial level in the following periods in which parity no longer helps moderate inflation. This contraction is basically caused by the demand inertia following upon the initial appreciation (Chart 6), not via interest rates (Chart 5), since economic agents very quickly anticipate that they will fall in long-term nominal values. The comparison with simulation 1 clearly shows that inflation responds more quickly to a temporary increase in interest rates, although the convergence towards the long-term inflation rate target of 3% is faster if the reduction in interest rates is immediate. This apparent asymmetry is explained almost completely by the different pattern of response of the nominal exchange rate and it confirms the predominance of the competitiveness channel in the monetary transmission mechanism in the case of Spain.

In the exercise shown in simulation 1, the initial depreciation had an inflationary impact; later on, as the nominal exchange rate stabilises, a swift drop in competitiveness occurs,[33] which helps dampen inflation. When, on the contrary, the temporary increase in interest rates causes the exchange rate to over-react (simulation 2), the strong initial appreciation has an immediate deflationary impact. However, after the initial phase, and while the interest rates differential with the outside world continues to be positive (and even higher than in the baseline), the national currency depreciates and part of the competitiveness lost is recovered. This helps cushion the impact of recession and induces a slight rise in inflation.

Even when the monetary authorities opt for a temporary increase in interest rates to induce a rapid drop in inflation, it definitely seems that the best thing to do is to shorten the period during which an interest rate differential causes a temporary depreciation (see Chart 6). If the central bank bears full credibility on its efforts to reduce inflation, then it will be able to proceed quickly to lower interest rates in a manner compatible with that objective. One way of increasing credibility is to carry out supply-side structural reforms.[34] Another way would be to obtain fiscal policy backing. This paper does not, however, go into such coordinated strategies.

Given that the assumption in the simulations regarding the initial response of the currency remains arbitrary, two further simulations were carried out in order to compare the sensitivity of the findings to the behaviour of the exchange rate. On the assumption that there is an initial variation in the nominal interest rate analogous to that in simulation 2, the nominal exchange rate is allowed to appreciate *immediately by 0%* (simulation 3) and *10%* (simulation 4), respectively. Following this initial reaction, there is a subsequent depreciation of 8 points derived from the accumulated interest rates differential. The behaviour of the rate of inflation and the rate of growth of GDP are also shown in Charts 7 and 8, along with the results of simulation 2 (appreciation of 5 percent). With an effective 10% appreciation of the peseta, we note that inflation declines by more than two points in the first year, in part due to a slight 0.3% recession compared to growth at the outset, but mainly due to the rapid decline in external inflation and its favourable impact on domestic price formation. On the other hand, if we assume no appreciation, monetary policy induces a path in which inflation declines, but more slowly.

The short and long-term output costs of the various monetary policy strategies are summarised in Table 3. The long-term costs are obtained by adding the GDP losses incurred each year. Having constructed the model with GDP growth rates, our measure for assessing costs is given

---

[33] Chart 6 shows an appreciation of the real exchange rate from the first period onwards.

[34] Which the model could take into account by altering the constant term in equation (6), for instance, or via an autonomous reduction in inflationary expectations.

by the GDP level itself, which does not necessarily have to return to the original level.[35] By this criterion, the long-term cost of the permanent 2-point drop in inflation is the same in all the simulations: a little under 1% of the original GDP[36] per year. The results in the table also show, in each simulation, the importance of short-term behaviour as the key factor to be borne in mind in assessing the relative advantages of each kind of monetary policy. The sacrifice ratio (growth in output divided by the change in inflation) in the first two years of simulations 2, 3, and 4 indicates that the lower the exchange rate appreciation associated with the temporary increase in interest rates, the lower the cost of the policies in those years.

Table 3

Impact and real cost of the simulations

Simulation	Initial change in the exchange rate[1]	Inflation differential				Growth of output differential				
		year 1	year 2	year 3	year 4	year 1	year 2	year 2	year 4	long term[2]
Monetary policies										
1. Immediate reduction of interest rates with initial depreciation	5%	1.24	0.29	-0.16	-0.77	0.20	0.01	-0.01	-0.22	-0.92
2. Temporary increase in interest rates and initial appreciation	-5%	-1.41	-0.87	-0.90	-0.86	-0.25	-0.15	-0.13	0.04	-0.92
3. Temporary increase in interest rates	0%	-0.07	0.15	-0.02	-0.24	-0.01	0.00	0.00	-0.01	-0.92
4. Temporary increase in interest rates and initial appreciation	-10%	-2.79	-1.93	-1.82	-1.51	-1.51	-0.31	-0.27	0.08	-0.92

[1] A positive figure denotes a depreciation. [2] Accumulated sum of the growth in output once the model reaches the steady-state solution.

[35] As was argued earlier on, a transitory response of the rate of growth to deflationary policy assumes that there is a permanent effect on the level of output. In that sense, the model is not strictly neutral, even though it returns to the initial rate of growth of GDP.

[36] This drop in GDP is the result of a corresponding 5% loss of competitiveness.

Conclusions

We have focused on an aggregate model of the monetary transmission mechanism in the Spanish economy, designed to evaluate, by simulations, alternative monetary policy strategies. The model incorporates the main channels through which central bank policy can affect the inflation rate. The parameters for the model were based on data for the 1973-94 period. Nonetheless, the model was specifically construed to reflect the main features of monetary transmission today. Thus, monetary policy is defined as a short-term interest rate path geared to achieving an inflation target. At the same time, the increasing openness of the economy reinforces the role of competitiveness in the transmission of monetary impulses. Finally, the efficiency of financial markets means that the impact of economic agents' expectations regarding the future behaviour of asset prices must be taken explicitly into account.

Solving the model has drawn attention to a few consequences of abandoning the establishment of an intermediate objective in terms of some monetary aggregate. Automatic adjustment of the money supply, aimed at keeping the interest rate within established values, leads to the indeterminacy of nominal magnitudes, particularly the price level and interest rate. Although it is true that this indeterminacy has no impact on the value of real variables, the short-term responses of inflation and competitiveness are indeed affected. To sidestep this problem, the model imposes a terminal exchange rate value, which has a considerable effect on the short-term cost of the different disinflationary policies. With due caveats from this limitation, the simulations do show a series of regularities, which are summarised very briefly in Table 3.

The achievement of a permanent reduction of the Spanish rate of inflation to the average level of inflation prevailing in Spain's trading partners presupposes, sooner or later, reducing nominal interest rates by at least as much as the current inflation differential. A policy in which that reduction in nominal interest rates is carried out immediately runs the risk of generating a short-term resurgence of inflation, unless it is accompanied by tougher fiscal policy or appropriate supply-side policies. If these policies are not promptly and forcefully applied, the currency depreciates and this leads to a period of temporarily higher inflation. That, in turn, could seriously erode the anti-inflationary credibility of the monetary policy strategy being pursued, making the ultimate objective harder to attain.

Judging by the simulations carried out, the most credible way to reduce inflation is by temporarily increasing interest rates, which leads to a temporary slowing down of the rate of output growth. In this case, inflation converges towards its long-term value, remaining below the initial figure, while the magnitude of the initial response depends on the degree of appreciation of the nominal exchange rate. The simulations also show that the degree of competitiveness plays a very important part in monetary transmission. Moreover, this effect acts powerfully on the supply side, lowering expectations of imported inflation and, with them, wage demands.

Although the long-term effects of all the anti-inflationary strategies analysed are similar, the structure over time of the response of inflation and GDP depends crucially on how competitiveness develops. Indeed, variations in the real exchange rate appear to be the most important of the channels making up monetary transmission mechanism in our country. The fact that tradable goods become relatively more expensive as a result of a more restrictive monetary policy induces a fall in aggregate demand and dampens inflationary expectations. The magnitude of this expectations or supply effect is one of the key determinants of the strength of the transmission mechanism and more empirical evidence has to be gathered concerning it. Likewise, it is necessary to broaden our knowledge of the response of the nominal exchange rate to variations in the expected path of interest rates in order to overcome one of the main limitations in the study of the monetary transmission mechanism within a framework of direct inflation targets.

Appendix A: The estimated model

The sample period is 1970Q1 - 1994Q4. All the equations have been estimated using instrumental variables such as the lags of the variables up to the eighth period as well as deterministic variables. The figures in brackets indicate the t-statistical for regressions and, for autocorrelation tests, the degree of marginal significance.

Equation IS (1):

$$\Delta y_t = \underset{(0.50)}{0.001} + \underset{(16.77)}{1.1489}\Delta y_{t-1} - \underset{(2.49)}{0.2610}\Delta y_{t-2} + \underset{(1.80)}{0.1498}\Delta y_{t-3} - \underset{(3.79)}{0.2502}\Delta y_{t-4} + \underset{(4.83)}{0.0706}\Delta\left(\frac{m}{p}\right)_{t-1}$$

$$+\underset{(3.31)}{0.0521}\Delta^2\left(\frac{m}{p}\right)_{t-6} + \underset{(1.69)}{0.0506}\Delta y^*_{t-2} + \underset{(3.41)}{0.0887}\Delta y^*_{t-4} - \underset{(1.82)}{0.0390}\Delta y^*_{t-8} + \underset{(4.09)}{0.0232}\Delta q_t - \underset{(3.87)}{0.0134}\Delta^2 q_{t-6}$$

$$+\underset{(2.76)}{0.0462}\Delta r_{t-1} - \underset{(3.58)}{0.0631}\Delta r_{t-7} + \underset{(6.87)}{0.2441}\Delta^2 g_t - \underset{(3.11)}{0.1089}\Delta^2 g_{t-1} + \underset{(3.57)}{0.1076}\Delta^2 g_{t-2} + \underset{(3.57)}{0.1076}\Delta^2 g_{t-7}$$

$$+\underset{(3.65)}{0.0042}\,dummy\ 80Q1 - \underset{(13.04)}{0.0165}\,dummy\ 80Q2 - \underset{(2.20)}{0.0025}\,dummy\ 92Q3$$

$\overline{R}^2 = 0.95$; standard error = 0.0010; first-order autocorrelation test:

$\underset{(0.09)}{X^2(1) = 2.7099}$; first to fourth-order autocorrelation test: $\underset{(0.41)}{X^2(4) = 3.927}$

Money demand equation LM (2):

$$\Delta\left(\frac{m}{p}\right)_t = \underset{(4.66)}{-2.2652} + \underset{(6.66)}{0.5519}\Delta\left(\frac{m}{p}\right)_{t-1} - \underset{(1.60)}{0.1953}\Delta R_t + \underset{(2.54)}{0.2746}R_{t-1} - \underset{(2.29)}{0.2265}\Delta R_{t-2} - \underset{(1.62)}{0.0551}\Delta i_{t-3}$$

$$-\underset{(3.11)}{0.4256}\Delta^2 p_t - \underset{(5.33)}{0.406}\left(\frac{m}{p}\right)_{t-1} + \underset{(5.06)}{0.5105}y_{t-1} + \underset{(5.13)}{0.0016}\,trend - \underset{(5.16)}{0.0163}\,seasonal\,T1$$

$$-\underset{(2.34)}{0.0067}\,seasonal\,T2 - \underset{(2.30)}{0.0117}\,step92Q1 + \underset{(2.06)}{0.0111}\,step93Q4$$

$\overline{R}^2 = 0.72$; standard error = 0.0047; first-order autocorrelation test:

$\underset{(0.12)}{X^2(1) = 2.3647}$; first to fourth-order autocorrelation test: $\underset{(0.004)}{X^2(4) = 15.02}$

Interest rates term structure equation TSIR (4):

$$(R_t - i_t) = 0.91(R_{t+1} - i_t)$$

(13.3)

$\overline{R}^2 = 0.72$; standard error $= 0.8731$; first-order autocorrelation test:

$X^2(1) = 11.10$; first to fourth-order autocorrelation test: $X^2(4) = 14.41$

(0.0008) (0.006)

Aggregate supply equation AS (6):

$$\Delta p_t^d = 0.8061\Delta p_{t-1}^d + 0.1563\Delta P_{T-7}^d - 0.2003\Delta^2 p_{t-1}^d + 0.0241(\Delta p_{t-1}^* + \Delta e_{t-1}) + 0.0185(\Delta^2 p_{t-5}^* + \Delta^2 e_{t-5})$$

(14.85) (3.05) (1.96) (1.56) (1.52)

$$+0.26.7\Delta_{t-1} - 0.0021\, seasonal\, T1 - 0.0014\, seasonal\, T2 - 0.0021\, seasonal\, T3$$

(3.73) (1.65) (1.19) (1.72)

$\overline{R}^2 = 0.87$; standard error $= 0.0045$; first-order autocorrelation test:

$X^2(1) = -0.09396$; first to fourth-order autocorrelation test: $X^2(4) = 2.6964$

(0.60)

CPI equation (7):

$$(\Delta p_t - \Delta p_t^d) = -0.1453\Delta^2 p_{t-2} - 0.1598\Delta^2 p_{t-3} + 0.1481\Delta^2 p_{t-6} + 0.1823\Delta^2 p_{t-8} - 0.6957\Delta^2 p_t^d$$

(1.85) (2.16) (2.30) (2.85) (2.62)

$$+0.0497\Delta^2 p_t^* + 0.9458\Delta^2 p_{t-1}^*$$

(1.30) (1.70)

$\overline{R}^2 = 0.29$; standard error $= 0.0045$; first-order autocorrelation test:

$X^2(1) = 3.4001$; first to fourth-order autocorrelation test: $X^2(4) = 6.09$

(0.06) (0.19)

Appendix B: Model dynamics

Let us consider the following model of a small open economy, simpler than the one estimated but containing the same basic properties:

$$y = \delta\left(e - p + p^{*}\right) - \gamma\left(r - \dot{p}\right) \tag{B.1}$$

$$\dot{p} = \pi + \phi y \tag{B.2}$$

$$\dot{\pi} = \beta\left(\dot{p} - \pi\right) \tag{B.3}$$

$$\dot{e} = r - r^{*} \tag{B.4}$$

$$m - p = \kappa y - \lambda r \tag{B.5}$$

The log of output (y) is a decreasing function of the (ex-post) real interest rate $(r - \dot{p})$ and an increasing function of the log of real exchange rate ($e - p + p^*$). Price inflation depends on expectations (π) and the output gap. The log of the demand of real money balances (m-p) is a function of income and the nominal interest rate. Inflation expectations $(\dot{\pi})$ react to past forecast errors. We also assume that the uncovered parity condition holds and that there is perfect foresight in the foreign exchange market, as widely used in many open economy models.

When the central bank follows an interest rule (setting r to r^0) the money supply is endogenous, and the economy is left without a nominal anchor. The state space representation of the system (B.1) - (B.5) can be written as in (B.6),

$$\begin{bmatrix} \dot{p} \\ \dot{\pi} \\ \dot{e} \end{bmatrix} = \begin{bmatrix} \frac{-\phi\delta}{1-\phi\gamma} & \frac{1}{1-\phi\gamma} & \frac{\phi\delta}{1-\phi\gamma} \\ \frac{-\beta\phi\delta}{1-\phi\gamma} & \frac{-\beta\phi\gamma}{1-\phi\gamma} & \frac{\beta\phi\delta}{1-\phi\gamma} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p \\ \pi \\ e \end{bmatrix} + \begin{bmatrix} \frac{\phi\delta}{1-\phi\gamma} & 0 & \frac{-\phi\gamma}{1-\phi\gamma} \\ \frac{\beta\phi\gamma}{1-\phi\gamma} & 0 & \frac{-\beta\phi\delta}{1-\phi\gamma} \\ 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} p^{*} \\ r^{*} \\ r^{0} \end{bmatrix} \tag{B.6}$$

Since the determinant is zero, the system has one zero root and the long-run value of the endogenous variables remains indeterminate. A closer look at the dynamic system reveals the cause of this indeterminacy. According to the uncovered parity condition, the exchange rate has a unit root that cannot be removed; hence, the bottom row of the first matrix in (B.6) is a row of zeroes. Solving the third row of the system (B.6) as an arbitrage condition gives the following expression for the exchange rate:

$$e(t) = -\int_{t}^{T}\left(r^{0}(s) - r^{*}(s)\right)ds + e^{e}(T) \tag{B.6'}$$

The nominal exchange rate is a non-predetermined or forward-looking variable, whose value at a particular point in time cannot be tied down by its initial condition. The money supply is endogenous and so is its value at T ($m(T)$). Hence the terminal value $e(T)$ is not independent of the particular path the economy follows and cannot be used to fix the current value $e(t)$. In other words, under rational expectations, the expected interest rate differential fixes the slope (the rate of change) of the nominal exchange rate, but not its level.

The fixed interest rate rule implicit in the model is perhaps too extreme, but it should be clear that the problem of indeterminacy would arise whatever the monetary rule we introduce as long as it is defined in real, rather than in nominal terms. A feedback rule like (B.7),

$$\dot{r} = \eta\left(r^0 - r\right) \tag{B.7}$$

such that the central bank smoothes its interest rate policy to achieve a target level (r^0), does not remove the zero root in the dynamic system as can be seen in the state space representation (B.8):[37]

$$\begin{bmatrix} \dot{p} \\ \dot{e} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} \frac{-\phi\delta}{1-\phi\gamma} & \frac{\phi\delta}{1-\phi\gamma} & 0 \\ 0 & 0 & 1 \\ 0 & 0 & -\eta \end{bmatrix} \begin{bmatrix} p \\ e \\ r \end{bmatrix} + \ldots \tag{B.8}$$

It should be clear that despite the nominal indeterminacy problem, the model can still be used to study the long run. Although the nominal system is indeterminate, the real system is not. In fact, a change of variables in (B.6) and (B.8) would render the system well defined in terms of the real exchange rate (and inflation expectations). For example, the dynamic system in (B.6) could be rewritten as:

$$\left(\dot{e} - \dot{p}\right) = \frac{-\phi\delta}{1-\phi\gamma}\left(e - p\right) + \ldots$$

from which it is trivial to obtain the steady-state solution of the model:

$$\dot{e} = \dot{p} = \pi = r^0 - r^* \qquad y = \dot{\pi} = 0 \qquad r - \dot{p} = r^0 - \dot{p} = r^* - \dot{p}^* = r^*$$

Price level indeterminacy and disinflation

The problem of price indeterminacy under interest rate rules and rational expectations, is well known since the work of Sargent and Wallace (1975) and so are its solutions. It can be shown that any fixed or feedback rule defined in nominal terms would tie down a unique path for the nominal variables under the assumptions of convergence to the steady state. A monetary rule making *the level of the money supply* exogenous (Dornbusch (1976)) removes the source of indeterminacy. Similarly, the unit root in the nominal system might be removed if the central bank sets its interest rate rule so as to achieve a *target price level*. For instance, if the bank sets its interest rate according to the feedback rule (B.9), all the roots in the system (B.10) have non-zero real parts,

$$\dot{r} = \eta\left(p - p^0\right) \tag{B.9}$$

[37] Hereafter, the inflation expectations state variable (π) has been excluded from the dynamic system, since it is the interaction among the dynamics of p, e and r that causes the indeterminacy problem.

$$\begin{bmatrix} \dot{p} \\ \dot{e} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} \frac{-\phi\delta}{1-\phi\gamma} & \frac{\phi\delta}{1-\phi\gamma} & \frac{-\phi\gamma}{1-\phi\gamma} \\ 0 & 0 & 1 \\ \eta & 0 & 0 \end{bmatrix} \begin{bmatrix} p \\ e \\ r \end{bmatrix} + \ldots \qquad \text{(B.10)}$$

The target level of nominal money balances or of the price level become, in these settings, the nominal anchor that is lacking once the monetary authority moves to an interest rate rule. By the same token, the problem of nominal indeterminacy would disappear once a country has pegged its currency within a monetary union; the fixed *level of the exchange rate* would become the nominal anchor, pinning down the path of nominal prices at any point in time.

Nominal-targeting interest rules are not very common in central banking, however. Feedback rules are most often designed in terms of real variables, as in equation (B.7). The increasing concern about price stability has led most central banks to set the inflation rate at the heart of monetary policy making. Inflation-targeting has then become the most popular way of conducting the interest rate policy. A natural question arises at this point: does inflation-targeting get rid of the problem of nominal indeterminacy? The answer is that, under fairly general circumstances, it does not. Let us consider a feedback interest rule as in (B.11),

$$\dot{r} = \eta\left(\dot{p} - \dot{p}^0\right) \qquad \text{(B.11)}$$

Some authors have already pointed out the differences of price-targeting versus inflation-targeting in monetary policy. Kerr and King (1996) argue that, in a closed economy model even with an expectations-augmented IS schedule, a unique rational expectations equilibrium may be pinned down as long as the interest rate response to deviations of inflation from its target is "sufficiently aggressive" (i.e. if η is big enough). Unfortunately, a high η does not help to get a unique path for the nominal variables in a model with forward looking nominal variables. In our framework, it is easy to show that, despite its similarity with (B.9) an interest rate rule such as (B.11) does not help to remove the unit root from the dynamic nominal system in (B.12):

$$\begin{bmatrix} \dot{p} \\ \dot{e} \\ \dot{r} \end{bmatrix} = \begin{bmatrix} \frac{-\phi\delta}{1-\phi\gamma} & \frac{\phi\delta}{1-\phi\gamma} & 0 \\ 0 & 0 & 1 \\ \frac{-\eta\phi\delta}{1-\phi\gamma} & \frac{\eta\phi\delta}{1-\phi\gamma} & 0 \end{bmatrix} \begin{bmatrix} p \\ e \\ r \end{bmatrix} + \ldots \qquad \text{(B.12)}$$

It could be argued that the nominal exchange rate path might be unique even if the central bank sticks to a monetary policy rule like (B.11), when the economy is operating close to its inflation target $\left(\dot{p}^0\right)$. Since the central bank would only have to react to small temporary deviations of the inflation rate, it would seem fair to assume that $e^e(T)$ should not deviate too much from the value that was expected before the change in the interest rate was announced. When a country is well above its target inflation, matters are rather different. In this case, bringing inflation down requires a permanent change in the interest rate differential; i.e. a permanent change in the rate of depreciation. Following this, it is no longer justified to assume that the steady state value of $e(e^e(T))$ remains unchanged. In such a case, unless the monetary authority sets a reaction function in terms of some nominal target, the initial reaction of the nominal exchange rate cannot be signed. Furthermore, even if the financial markets would interpret the movement in the interest rate as a tightening in the monetary policy, so that we could expect an instantaneous appreciation of the currency, the model in (B.1) - (B.5) would still be silent about the size of this loss of competitiveness and hence about the path of inflation and prices.

Stability of the model

We analyse the stability conditions of the real system, written as (B.13) for convenience.

$$\begin{bmatrix} \dot{e}-\dot{p} \\ \dot{\pi} \end{bmatrix} = \begin{bmatrix} \dfrac{-\phi\delta}{1-\phi\gamma} & -\dfrac{1}{1-\phi\gamma} \\ \dfrac{\beta\phi\delta}{1-\phi\gamma} & \dfrac{\beta\phi\gamma}{1-\phi\gamma} \end{bmatrix} \begin{bmatrix} e-p \\ \pi \end{bmatrix} + \begin{bmatrix} \dfrac{-\phi\delta}{1-\phi\gamma} & -1 & \dfrac{1}{1-\phi\gamma} \\ \dfrac{\beta\phi\delta}{1-\phi\gamma} & 0 & \dfrac{-\beta\phi\gamma}{1-\phi\gamma} \end{bmatrix} \begin{bmatrix} p^* \\ r^* \\ r^0 \end{bmatrix} \tag{B.13}$$

The determinant of the matrix is positive, so that the two roots (solutions of (B.14)) are equally signed. Since one of the variables is non-predetermined (*e-p*) and the other is predetermined (π), the transversality condition is not sufficient to pin down a unique dynamic path.

$$\lambda = \frac{1}{2}\left[\frac{\phi(\beta\gamma-\delta)}{1-\phi\gamma} \pm \left[\left(\frac{\phi(\beta\gamma-\delta)}{1-\phi\gamma} \right)^2 - 4\frac{\beta\phi\gamma}{1-\phi\gamma} \right]^{\frac{1}{2}} \right] \tag{B.14}$$

The system can be either globally stable or globally unstable, depending on the sign of $(\beta\gamma - \delta)$. When the system is stable $(\beta\gamma < \delta)$, setting the interest rates suddenly and permanently to the international level is a feasible deflationary policy. This is the assumption that was made in the main text. Furthermore, without loss of generality, it has also been assumed that the roots have imaginary parts (i.e. in B.14, the square root term is negative).

Conversely, an unstable model will rule out a sudden and permanent drop in the interest rate to the international level. Rather, the only possible monetary policy will consist of an increase in the interest rate to bring inflation down, followed by a permanent fine-tuning of the policy to keep inflation within bounds.

References

Alvarez, L. J., F. C. Ballabriga and J. Jareño (1995): "Un modelo macroeconométrico trimestral para la economía española". Banco de España, *Working Paper*, No. 9524.

Ayuso, J., A. Haldane and F. Restoy (1994): "Volatility transmission along the money market yield curve". Banco de España, *Working Paper*, No. 9403.

Andrés, J., A. Escribano, C. Molinas and D. Taguas (1989): "La inversión en España". *Moneda y Crédito*, pp. 67-95.

Ayuso, J. and S. Núñez (1997): "The yield curve", in *Monetary Policy and Inflation in Spain*, Banco de España, forthcoming.

Ballabriga, F. C. and M. Sebastián (1993): "Déficit Público y tipos de interés en la economía española: ¿existe evidencia de causalidad?" *Revista Española de Economía*, 16, pp. 283-306.

Bajo, O. and M. Montero (1995): "Un modelo ampliado del comercio exterior español". *Moneda y Crédito*, 201, pp. 153-182.

Bernanke, B. and M. Gertler (1995): "Inside the black box: the credit channel of monetary policy transmission". *Journal of Economic Perspectives*, 4, pp. 27-48.

Buisán, A. and E. Gordo (1994): "Funciones de importación y exportación de la economía española". *Investigaciones Económicas*, 18, pp. 165-192.

Buiter, W. and M. Miller (1981): "Monetary Policy and International Competitiveness: the problems of adjustment". *Oxford Economic Papers*, pp. 143-175.

Cabrero, A., J. L. Escrivá and T. Sastre (1993): "Demand equations on the new monetary aggregates". *Serie de Estudios Económicos*, No. 52, Banco de España.

Canzoneri, M. and B. Diba (1997): "Fiscal restrictions on Central Banks' Independence and on Price Stability", in *Monetary Policy and Inflation in Spain*, Banco de España, forthcoming.

Dhar, J. K., P. G. Fisher, A. M. Holland and D. L. Pain (1995): "Interest rate effects in the Bank of England medium-term forecasting model" in *Financial structure and the monetary policy transmission mechanism*, Bank for International Settlements, pp. 537-542.

Dornbusch, R. (1976): "Expectations and Exchange rate dynamics". *Journal of Political Economy*, No. 84, pp. 1161-1176.

Duguay, P. (1994): "Empirical evidence on the strength of the monetary transmission mechanism in Canada". *Journal of Monetary Economics*, 33, pp. 39-61.

Escrivá, J. L. and R. Santos (1991): "Un estudio del cambio de régimen en la variable instrumental del control monetario en España". Banco de España, *Documento de Trabajo*, No. 9111.

Estrada, A., I. Hernando and J. Vallés (1997): "The impact of interest rates on private expenditure", in *Monetary Policy and Inflation in Spain*, Banco de España, forthcoming.

Estrada, A., T. Sastre and J. L. Vega (1994): "El mecanismo de transmisión de los tipos de interés: el caso español". Banco de España, *Documento de Trabajo*, No. 9408.

Estrada, A. (1993): "Una función de consumo de bienes duraderos". *Revista Española de Economía*, 10, pp. 135-162.

Haldane, A. and C. K. Salmon (1995): "Three issues on inflation targets", in A. Haldane (ed.) *Targeting Inflation*, Bank of England.

Hernando, I. (1997): "The credit channel in monetary policy transmission", in *Monetary Policy and Inflation in Spain*, Banco de España, forthcoming.

Marín, J. and J. M. Peñalosa (1997): "The institutional framework and budgetary policy: implications for monetary policy in Spain", in *Monetary Policy and Inflation in Spain*, Banco de España, forthcoming.

McCallum, B. T. (1981): "Price level determinancy with interest rate policy rule and rational expectations". *Journal of Monetary Economics*, 8, pp. 319-329.

Meltzer, A. (1995): "Monetary, Credit and (other) transmission processes: A monetarist perspective". *Journal of Economic Perspectives*, 4, pp. 49-72.

Poole, W. (1970): "Optimal choice of monetary policy instruments in a simple stochastic macro model". *Quarterly Journal of Economics*, 84, pp. 197-216.

Raymond, J. L. and J. Palet (1990): "Factores determinantes de los tipos de interés reales en España". *Papeles de Economía Española*, 43, pp. 144-160.

Roubini, A and S. Kim (1995): "Monetary Policy and Exchange rates in Europe". Yale University, mimeo.

Sargent, T. and N. Wallace (1975): "Rational expectations, the optimal monetary instrument and the optimal money supply rule". *Journal of Political Economy*, 83, pp. 241-254.

Taylor, J. (1993): "*Macroeconomic Policy in a world economy*". W. W. Norton (ed.).

Taylor, J. B. (1995): "The monetary transmission mechanism: an empirical framework". *Journal of Economics Perspectives*, 4, pp. 11-26.

Vega, J. L. (1997): "The ALP long-run demand function" in *Monetary Policy and Inflation in Spain*, Banco de España, forthcoming.

Comments on: "A structural model for the analysis of the impact of monetary policy on output and inflation" by Javier Andrés, Ricardo Mestre and Javier Vallés

by Palle Andersen

The main purpose of this paper is to simulate the effects of various changes in monetary policy, using an aggregate macroeconomic model for Spain. The first two sub-sections review the various transmission channels of monetary policy, starting with a change in a monetary aggregate and then turning to changes in interest rates. In this respect three principal differences are highlighted: (i) a *permanent* reduction in monetary aggregate growth is equivalent to a *temporary* change in interest rates; (ii) to emulate the effects of a *reduction* in monetary growth, the equivalent change in interest rates requires an initial *increase*, followed by a reduction later on; and (iii) contrary to an intermediate aggregate target, a framework based on interest rate control lacks a nominal anchor. This is, of course, true but, as discussed below, I would not consider this to be such a serious problem.

The paper then presents the aggregate model, with actual estimates given in Annex A and discussed in Section 1.1. I have some problems with this part. The discussion of wealth effects is not very clear and it is hard to understand how a net debtor country like Spain experiences a rise in net wealth when the currency depreciates. The arbitrage condition applied in estimating the yield curve is also hard to understand and the actual estimate looks rather implausible. Finally, it seems rather odd to have import prices influencing domestic inflation (the GDP deflator) and then estimating the CPI equation using a cointegration relation.

These are, however, minor concerns. My major problems with the paper lie elsewhere:

1. As already indicated, I do not share the authors' concern about the lack of a nominal anchor when the central bank is targeting a nominal exchange rate or the rate inflation. It is, of course, true that an inflation rate target leaves the price level undetermined but I am tempted to say: "So what?". When inflation is as low as it now is in Spain it is hard to believe that anybody would be much concerned about the price level. Moreover, I do not quite understand the authors' treatment of the central bank reaction function. While they recognise that a reaction function would create the nominal anchor, they rule it out on the grounds that this is not the way that Bank of Spain operates. I do not dispute that, but if the Bank has an inflation or exchange rate target, it surely must be reacting to *something* and that *something* must be the equivalent of a reaction function. Finally, the authors seem to slightly misrepresent the policy relevance of a nominal anchor. When a central bank targets the rate of inflation, it essential attempts to anchor expectations to the inflation target. In other words, with such a policy I would view the rate of inflation and not the price level as the policy relevant nominal anchor.[1]

2. My second major problem concerns the conclusion about the transmission of policy changes via movements in the exchange rate. When stating the conclusion, the authors seem to forget that exchange rate response was *imposed* and not the estimated result of a gradual opening of the Spanish economy. Moreover, their assumption that a 2 percentage point change in the interest rate generates a 5 percentage point change in the exchange rate seems *very* high. Would it not have been possible to estimate an exchange rate equation (other economists at the Bank of Spain have been quite successful) and get some solid ground for this conclusion?

---

[1] It is, of course, true that the rate of inflation is made stationary through targeting, the price level could be I(1). However, it might be recalled that with an aggregate money supply growth target, the money stock is equally likely to be I(1) ("base drift"). Moreover, if a central bank were to target the price level rather than the rate of inflation, there is a high risk of a very volatile monetary policy.

3. My third problem is perhaps a minor one but concerns the simulations based on an initial *reduction* in interest rates as an anti-inflationary strategy. If the ultimate aim is to reduce inflation, an action which initially generates higher inflation, is, as the authors themselves say, unlikely to be regarded as credible. So why have it? I think the paper would have gained by concentrating on the effects of relevant and credible policy changes.

4. Finally, I find an implied sacrifice ratio of only 0.5 much too low for a country like Spain with well known and large labour market rigidities. Given the aggregate nature of the model, labour market rigidities have not been included but for the next version of the model it might be a good idea to extend it to include the labour market and thus enable it to generate more plausible and credible sacrifice ratios.

In short, I regard the paper as a valid attempt to provide simulations in a transparent way for which the authors deserve credit. However, the paper suffers from some self-inflicted problems and overly strong assumptions.

The lags of monetary policy

David Gruen, John Romalis and Naveen Chandra[1]

Introduction

Good policy-making requires an appreciation of the dynamic relationship between the monetary policy instrument – the overnight cash rate – and the final objectives of policy – inflation and output. A thorough analysis of this relationship between instrument and objectives is a large task, however, because of the many transmission channels by which monetary policy influences the economy. Rather than examining each of these channels, this paper sets the more modest task of estimating the aggregate impact on Australian economic output of changes in the domestic short-term real interest rate.

The task of isolating the impact of monetary policy on output would be made much easier if we had a good explanation for the underlying business cycle. While this is obviously complex, we do have good empirical evidence over the past fifteen years that Australian output is strongly influenced by economic activity in the United States. US output is clearly "exogenous" to the Australian economy since it is not affected by either Australian output or Australian monetary policy. But the presence of this powerful exogenous influence makes it easier to identify econometrically the dynamic effect of monetary policy on Australian output.

Despite this econometric benefit, estimating the lags of monetary policy in Australia is not without its difficulties. As a rule, short-term real interest rates change only gradually, so that the current real interest rate is quite strongly correlated with interest rates in the recent past. As a consequence, it is hard to separate the effect on output of the current real interest rate from the delayed effects of the real rate in earlier periods. This problem leads to fairly wide margins of error in our estimates of the dynamic effect of monetary policy on output. Nevertheless, there is still strong evidence of an impact on output growth in the first, second and third years after a change in the domestic short-term real interest rate.

Another difficulty in isolating the dynamic impact of monetary policy on output arises from the forward-looking nature of policy. As well as responding to data about the past, policy-makers also act on information about current and future economic developments that is not part of any simple aggregate analysis of the relationship between monetary policy and output. This paper will show that this implies that standard estimation techniques underestimate the strength of monetary policy's impact on output, and overestimate the length of the lags of monetary policy.

The next section of the paper begins with an analytical discussion of the sources of the lags of monetary policy. It then turns to single equation models of Australian output which provide good empirical descriptions of the domestic business cycle over the past 15 years. The main focus of the section is to estimate the effect of a one percentage point change in the short-term real interest rate on output growth over the subsequent three years.

The following section, Section 2, discusses the implications of policy-makers responding to information that is not available to the econometrician estimating the relationship between

[1] Economic Group. The authors are grateful for comments from Adrian Pagan, colleagues at the Reserve Bank and participants at the January 1997 Meeting of Central Bank Model Builders and Econometricians at the Bank for International Settlements. They also thank Gordon Menzies for showing them his earlier work on instruments for the short-term real interest rate. The views expressed are those of the authors and should not be attributed to the Reserve Bank of Australia.

monetary policy and output. Under plausible assumptions, this is likely to result in an underestimation, using standard techniques, of the impact of monetary policy changes on output growth in the short run.

Making some allowance for this bias, we conclude that output growth falls by about one-third of one per cent in both the first and second years after a one percentage point rise in the short-term real interest rate and by about one-sixth of one per cent in the third year. Section 2 also finds that there is no evidence that the lags of monetary policy have become any shorter over the course of the 1990s.

The paper ends with a brief summary of the main results.

1. The lags of monetary policy

1.1 The sources of monetary policy lags

There are six main channels through which changes in interest rates affect economic activity: intertemporal substitution (since interest rates represent the relative price between the present and the future), the effect of induced changes in the exchange rate on the tradable sector, interest rate effects on other asset prices, cash-flow effects on liquidity constrained borrowers, credit supply effects, and the direct effect of changes in monetary policy on expectations of growth (Grenville 1996). Each of these channels – and the interaction between them – makes a contribution to the lags of monetary policy.

To begin at the beginning, however, the first source of monetary policy lags is the delay in the pass-through of changes in the overnight cash rate into other interest rates. While the response of short-term money market interest rates is rapid and complete, the pass-through to other interest rates, such as the deposit and lending rates of financial intermediaries, appears to be slower (Lowe 1995). Since intermediaries' interest rates are important determinants of cash-flow, asset prices, and the incentive to postpone expenditure, the slow pass-through contributes to the transmission lag from the real cash rate to activity.

Beyond the pass-through, an important source of lags arises from the gradual response of investment – both business investment and consumer investment in durables and dwellings – to changes in monetary policy. Adjustment costs associated with changing the level of the relevant capital stock are partly responsible. However, changes in interest rates also affect the incentive to postpone investment when returns are uncertain. The largely irreversible nature of many investments means that there is an option value to waiting to invest in a world of uncertainty (Dixit and Pindyck (1994)). When a firm or individual makes an irreversible investment, this option is exercised, eliminating the possibility of waiting for the arrival of new information that might have affected the timing or the desirability of the investment. A change in interest rates affects this option value, and will, therefore, affect the timing of the investment.

Empirical estimates for the United States suggest quite long lags in the adjustment of investment to shocks. For example, Jorgenson and Stephenson (1967) report a mean lag of seven quarters between changes in the rental price of capital and investment in US manufacturing, while Shapiro (1986) estimates that, in response to a shock to the required rate of return on capital, more than half the adjustment in the manufacturing capital stock occurs in the first year, but it takes over four years to be complete.

Turning to asset markets, economic theory would lead one to expect the full implications of a change in monetary policy to be incorporated rapidly into asset prices as soon as the change became apparent. In the important case of the exchange rate, however, this does not appear to occur. For example, Eichenbaum and Evans (1995) find, for the United States, that contractionary monetary policy leads to a prolonged gradual appreciation of the domestic currency with the maximal appreciation occurring after two to three-and-a-half years. As a consequence, the exchange rate effects on the tradable sector of the economy are also gradual and prolonged.

Finally, developments in one sector of the economy are gradually transmitted to other parts of the economy as agents, who were initially unaffected by the monetary policy change, respond to the altered behaviour of their suppliers and customers. These transmission channels to the wider economy also contribute to the aggregate lags of monetary policy.

1.2 Single equation models

Turning to empirical analysis, we begin with single equation models for Australian output. We use a general-to-specific modelling strategy in which insignificant lags of the variables are sequentially eliminated, leading eventually to parsimonious specifications. The models are variants of an earlier model estimated by Gruen and Shuetrim (1994).

We present two models which differ only in their treatment of inflationary expectations. After eliminating insignificant lags, both models take the form,

$$\Delta y_t = \alpha + \sum_{j=0}^{6}\beta_j r_{t-j} + \left[\gamma_2 \Delta f_{t-2} + \gamma_4 \Delta f_{t-4}\right] + \delta y_{t-1} + \chi w_{t-1} + \sum_{j=0}^{1}\phi_j \Delta w_{t-j} + \varepsilon_t \tag{1}$$

where Δy_t is quarterly growth of Australian non-farm output, r_t is the short-term real interest rate, Δf_t is growth of Australian farm output, y_{t-1} and w_{t-1} are the lagged log levels of Australian non-farm output and US output respectively, and ε_t is a mean-zero error term. Summary results for the two models, estimated by ordinary least squares, are shown in Table 1.[2]

The first two sets of independent variables model the influence of domestic variables on output. To control for domestic monetary policy, we use current and lagged values of the short-term real interest rate. With our focus on the length of the lags of monetary policy, we want to allow considerable flexibility in the estimated pattern of influence of monetary policy on output. We therefore use lags 0 to 6 of the short-term real interest rate, rather than eliminating all insignificant lags as we do for other variables.

We assume that inflationary expectations are backward-looking. For the underlying CPI model, we use the overnight cash rate set by the Reserve Bank minus underlying consumer price inflation over the past year to measure the short-term real interest rate, while for the headline CPI model, we subtract headline consumer price inflation over the past year.[3] For both models, the

---

2 We are faced with the common difficulty in econometrics that we require a timespan long enough to generate meaningful results but not so long that the underlying economic relationships change substantially during the estimation period. With this in mind, we omit the more financially regulated 1970s, and estimate from the financial year 1980/81 to the present;that is, 1980Q3 to 1996Q1, giving 63 quarterly observations. For our purposes, the float of the Australian dollar in December 1983 was not an important regime change because, from 1980 to 1983, the exchange rate was adjusted daily via a crawling peg with the US dollar and was, therefore, fairly flexible. For both models, the general specification from which we begin includes contemporaneous and four lags of farm output growth and US GDP growth as well as lags one to four of the dependent variable. A trend term is insignificant when added to either regression.

3 Rather than using past inflation, we could have generated estimates of the short-term real interest rate using a survey-based measure of inflationary expectations, although the only available measure of expected inflation in Australia spanning our estimation period is for consumers, whose expectations may not be representative of other economic agents (consumers' inflationary expectations are from the Melbourne Institute survey). Of the two measures of the past inflation used in our estimation, it is unclear which is a better measure of inflationary expectations in the economy. The headline measure is more widely reported but is directly affected by changes in the overnight cash rate (via their

coefficients on individual lags of the real interest rate are estimated imprecisely, but the mean of the real interest rate coefficients is negative, as expected, and highly significant (see Table 1).[4]

Table 1

Australian non-farm GDP growth regressions[1]

$$\Delta y_t = \alpha + \sum_{j=0}^{6} \beta_j r_{t-j} + \left[\gamma_2 \Delta f_{t-2} + \gamma_4 \Delta f_{t-4}\right] + \delta y_{t-1} + \chi w_{t-1} + \sum_{j=0}^{1} \phi_j \Delta w_{t-j} + \varepsilon_t$$

Variables	Underlying CPI model[2]	Headline CPI model[2]
Constant	24.64**	23.46**
	(2.82)	(2.64)
Real cash rate[3]	-0.035	-0.036
	{0.00}	{0.00}
Farm output, % change (lag 2)	0.020*	0.020*
	(2.39)	(2.26)
(lag 4)	-0.020*	-0.020*
	(-2.23)	(-2.04)
Lagged Australian GDP log level	-0.31**	-0.34**
	(-5.78)	(-5.96)
Lagged US GDP log level	0.38**	0.42**
	(6.02)	(6.21)
US GDP, % change[3]	0.047	0.061
	{0.00}	{0.00}
R^2	0.68	0.68
Adjusted R^2	0.60	0.59
Standard error of residuals	0.56	0.56
F-test for joint significance of Australian and US GDP levels	20.0	21.3
	{0.00}	{0.00}
LM test for autocorrelation of residuals:		
First order	1.49	0.083
	{0.22}	{0.77}
First fourth order	6.58	4.99
	{0.16}	{0.29}
Breusch-Pagan test for heteroscedasticity	20.27	16.83
	{0.09}	{0.21}

[1] The models are estimated by ordinary least squares using quarterly data over the period 1980Q3 to 1996Q1. Numbers in parentheses () are t-statistics. Numbers in brackets {} are p-values. Individual coefficients marked with * (**) are significantly different from zero at the 5% (1%) level. All variables in log levels and their differences are multiplied by 100 (so growth rates are in percentages). [2] To derive the real interest rate, inflation expectations are based on the underlying CPI or on the headline CPI. [3] The mean coefficient is reported for the real cash rate and US GDP % change to summarise the coefficients on these variables. The p-values are derived from F-tests of the joint significance of the lags.

effect on variable-rate housing mortgage interest rates); by contrast, the underlying measure, which excludes this direct effect, is a better measure of core consumer price inflation.

4 As a check of robustness, we also repeated the regression in the table using the yield gap (the cash rate minus the 10-year bond rate) instead of the real cash rate to control for the influence of monetary policy. The results are qualitatively similar, although both the explanatory power of the regression and the significance of this measure of monetary policy are much reduced.

The second set of domestic variables controls for the influence of farm output on the rest of the Australian domestic economy. Although the farm sector accounts for only about 4% of the Australian economy, widespread droughts, and the subsequent breaking of those droughts, lead to large changes in farm output which have multiplier effects on the wider economy.

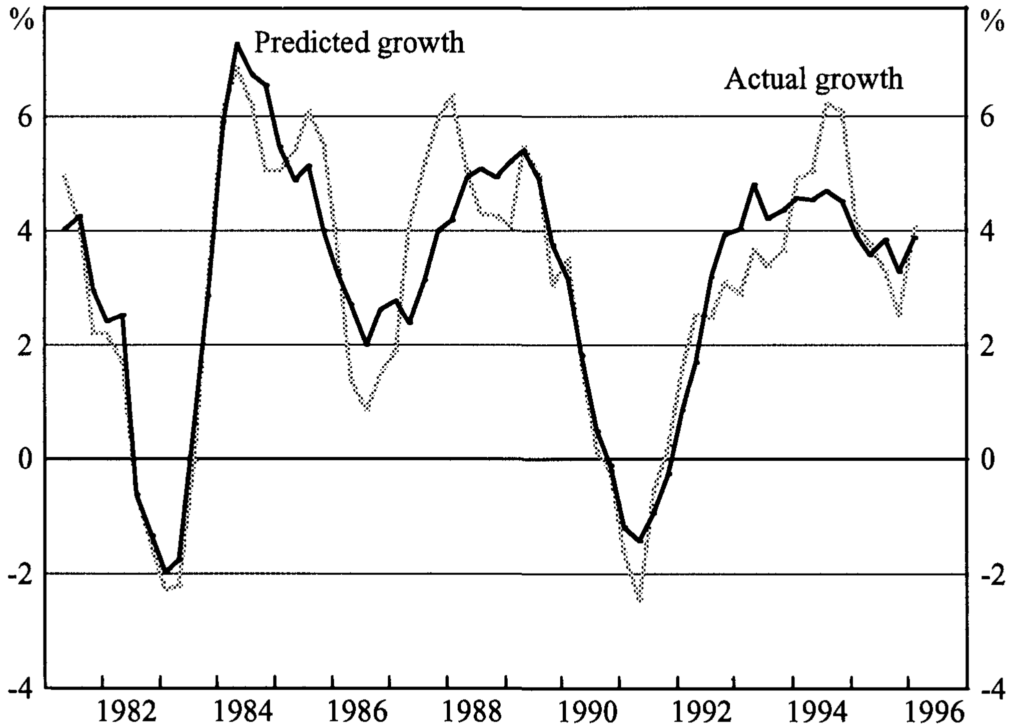
The rest of the independent variables in the equation control for both the short-run and longer-run effects of US output growth on Australian output. Including lagged log levels in the regression allows for a possible long-run (cointegrating) relationship between the log levels of Australian and US output, with the results providing strong evidence of the existence of this relationship.[5]

The importance of US output for the Australian business cycle, recently highlighted by McTaggart and Hall (1993), appears to arise for several reasons. In the shorter run, links between financial markets (Gruen and Shuetrim (1994); de Roos and Russell (1996); and Kortian and O'Regan 1996)), effects on Australian business confidence (Debelle and Preston (1995)) and a disproportionately large response of Australian exports to the US business cycle (de Roos and Russell (1996)) all play a role. In the longer run, technology transfer from the US seems to be important (de Brouwer and Romalis (1996)).

Figure 1

Actual and predicted Australian non-farm GDP growth

Four-quarter-ended growth, underlying CPI model



---

[5] Augmented Dickey-Fuller tests do not reject the hypothesis that the log levels of GDP are stochastically non-stationary I(1) variables (Gruen and Shuetrim (1994)). In the regressions in Table 1, the F-statistic for the joint significance of the lagged log levels of Australian and US GDP can be used to test for the existence of a long-run relationship between these variables and whether they are I(0) or I(1) variables. It does, however, have a non-standard distribution. For both regressions, we reject the null of no long-run relationship at the 1% level based on critical values tabulated in Pesaran, Shin and Smith (1996).

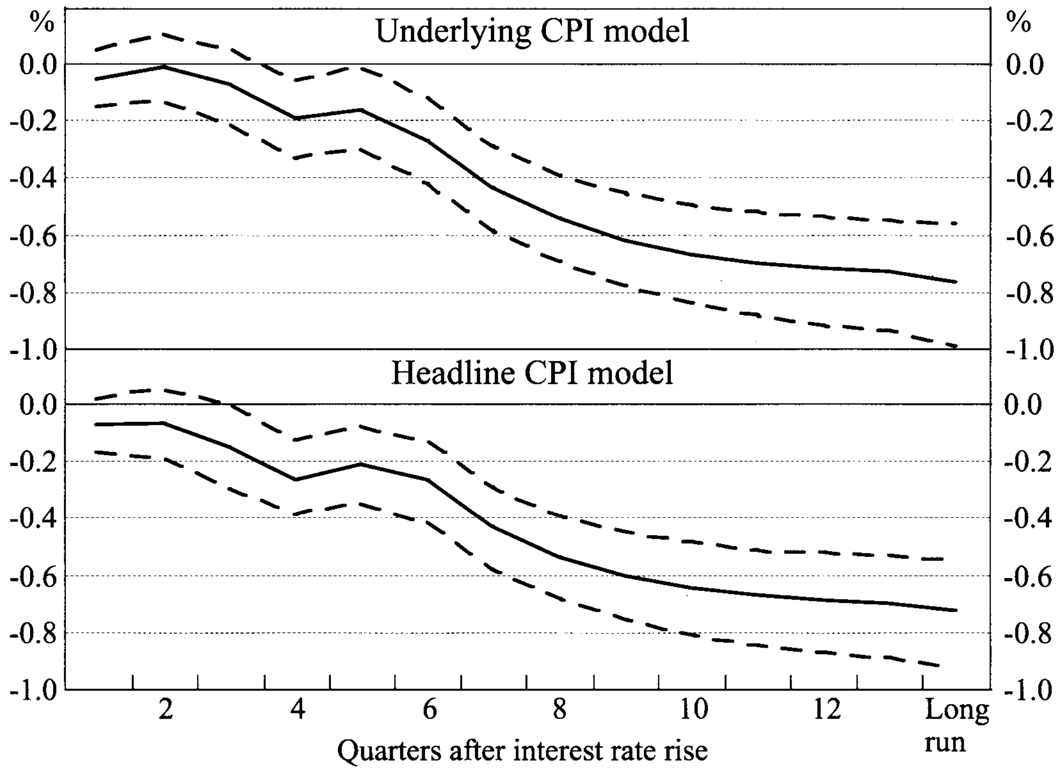
Both model regressions have appealing statistical properties, with no evidence of first-order or first to fourth-order serial correlation and no strong signs of heteroscedasticity. Despite their simplicity, the equations explain a substantial part of the variation in Australian quarterly non-farm GDP growth, with adjusted R^2s of 0.60 and 0.59. Both models explain the major features of the Australian business cycle since 1980. Figure 1 shows the results from the underlying CPI model.

1.3 Quantifying the lags of monetary policy

We now return to the lags of monetary policy, and examine the impact on domestic output of a sustained one percentage point rise in the domestic short-term real interest rate. Of course, in conducting this exercise, we should not lose sight of the fact that the domestic real rate is determined in the longer run by the world real rate rather than by domestic monetary policy.[6]

Figure 2

Impact of monetary policy on level of output



Note: The figure shows point estimates and 90% confidence intervals for the impact on the level of non-farm output of a one percentage point rise in the short-term real interest rate at the beginning of quarter 1.

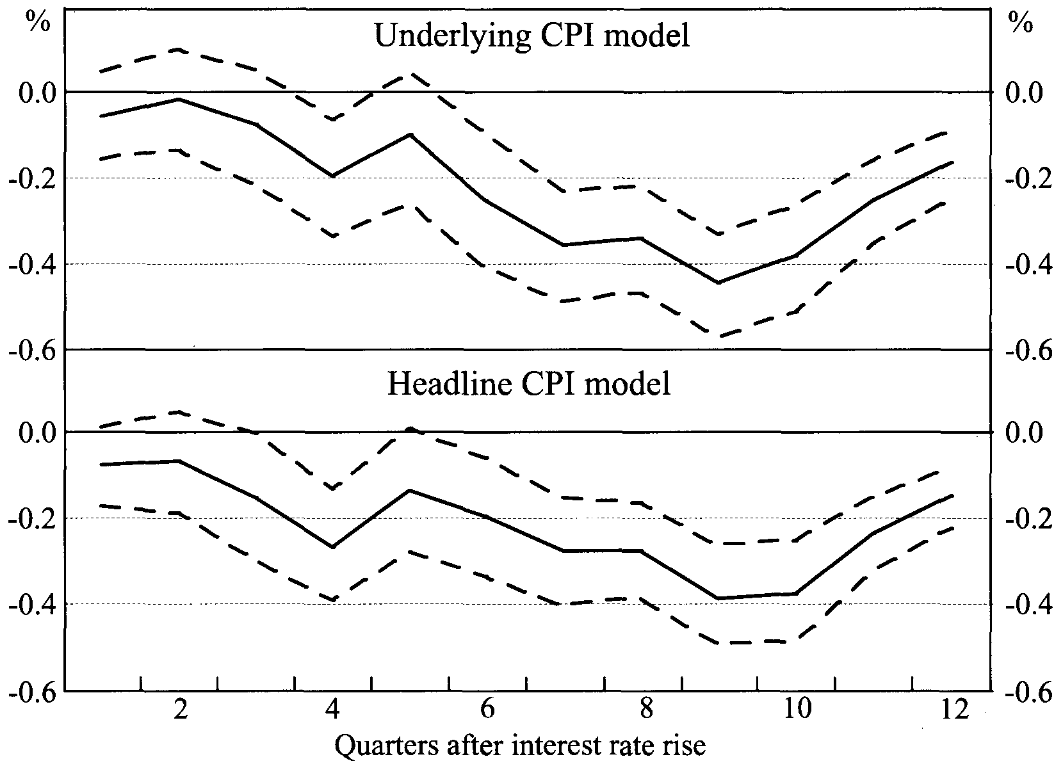
For both models, Figure 2 shows the effect of the rise in the domestic real interest rate on the level of non-farm output, while Figure 3 shows the effect on the year-ended growth of non-farm output. Both figures show point estimates and 90% confidence intervals. Since the effect on output of

6 Since Australia is small in the world capital market, the Australian short-term real interest rate is determined in the long run by the world short-term real interest rate plus or minus a risk premium (assuming no long-run trend in the Australian real exchange rate). Nevertheless, domestic monetary policy determines the Australian short-term real interest rate for long enough to have an important influence on the Australian macroeconomy.

a rise in the real interest rate is a non-linear function of the model parameters, the confidence intervals are estimated using a Monte Carlo procedure described in Appendix A.[7]

Figure 3

Impact of monetary policy on four-quarter-ended growth



Note: The figure shows point estimates and 90% confidence intervals for the impact on four-quarter-ended growth of non-farm output of a one percentage point rise in the short-term real interest rate at the beginning of quarter 1.

The level of output in either model is estimated to fall slightly for the first few quarters after a rise in the short-term real interest rate, but the fall is statistically insignificant. Over time, however, the contractionary effect on output gets stronger and becomes increasingly significant. Almost all the effect on the level of output occurs within three years.

It is also clear, however, that the confidence intervals are rather wide. The current real interest rate and its lags are quite strongly correlated, leading to unavoidable problems of multicollinearity in the regressions.[8] As a consequence, it is hard to disentangle the effect on output of the current real interest rate from the delayed effects of the real rate in earlier quarters. In other words, it is hard to estimate accurately the length of the lags of monetary policy.

---

7 The rise in the real interest rate occurs at the beginning of quarter 1. Therefore, for quarters 1, 2 and 3, the effect on year-ended growth shown in Figure 3 is small, partly because the real interest rate has been raised for less than a year. The first year is defined to be from quarter 0 to 4, and the second year from quarter 4 to 8. The point estimates shown in the figures are median outcomes from the Monte Carlo simulations. As a consequence, they differ very slightly from results derived from the OLS regressions.

8 For example, the correlation coefficient between the current real interest rate, defined using underlying inflation, and its lags falls from 0.88 for the first lag to 0.74, 0.63, 0.50, 0.34 and 0.22 for the sixth lag.

Another way to highlight this problem is to compare results from the two models on the estimated effect on output growth in the first and second years after a rise in the real interest rate. Assuming inflationary expectations respond to underlying inflation, the fall in output growth is smaller in the first year than in the second (the point estimates are falls in growth of 0.20% and 0.34%), suggesting that the lags in the transmission of monetary policy to output are quite long. Alternatively, assuming inflationary expectations respond to headline inflation, the fall in output growth in the two years is almost the same (the point estimates are 0.26% and 0.28%), suggesting rather shorter lags. The point of this comparison is clear enough: subtle changes in assumptions about inflationary expectations can lead to somewhat different estimated results.

As might be expected, while the point estimates from the two models are different, these differences are not statistically significant. Although the underlying CPI model suggests that the contractionary impact is stronger in the second year, at conventional levels of significance we cannot reject the alternative hypothesis that the impact is in fact stronger in the first year.[9]

An alternative way to summarise the length of the lags of monetary policy is to calculate the average lag length, defined by

$$\text{Average lag length (in quarters)} = \sum_{i=1}^{\infty}(i-1)\Delta m_i \Big/ \sum_{i=1}^{\infty}\Delta m_i \qquad (2)$$

where Δm_i is the effect on non-farm output growth in quarter i of the one percentage point rise in the real interest rate.[10] Using this formula for the underlying CPI model, the average length of the monetary policy lag is 6.4 quarters; for the headline model, it is a slightly smaller 5.8 quarters. For the reasons explained above, these numbers are again estimated imprecisely, with the 90% confidence interval from 5.1 to 7.7 quarters for the underlying model, and 4.5 to 7.0 quarters for the headline model.

2. Taking account of the policy response

2.1 The problem with ordinary least squares estimation

Estimating the lags of monetary policy using the approach adopted in the last section has an important drawback. In setting monetary policy, policy-makers do not rely solely on data about the past, but base their decisions also on information about current and expected future developments in inflation and output. From an econometric viewpoint, this renders the real interest rate an endogenous variable, that is likely to be correlated with the current and future residuals in an output equation.

The nature of the problem can be explained using a simple model. Assume that output depends negatively on the real interest rate, while the real interest rate is set on the basis of incoming information about output. In symbols,

$$y_t = -r_t + u_t$$

$$r_t = I_t + v_t$$

---

[9] In about 12% of the Monte Carlo simulations of the underlying CPI model, the contractionary impact on output is stronger in the first year.

[10] To give an example, if the interest rate rise led to an immediate once-off fall in the level of output (and hence a once-off fall in output growth in quarter 1) the average lag length as calculated by equation (2) would be zero, as required.

where y_t is output, r_t is the real interest rate, u_t and v_t are independent error terms and I_t is information about current output and, therefore, correlated with current output ($E_t(I_t u_t) = \gamma$). The correlation γ is positive because monetary policy will be tighter when output is expected to be above average.

If the output equation, $y_t = \alpha r_t + u_t$, is now estimated by ordinary least squares (OLS) the resulting estimate of α is, on average, smaller in magnitude than the true value, $\alpha = -1$. If the variables u_t, v_t and I_t have zero means and unit variances, the ratio R_α of the OLS estimate of α to its true value is $R_\alpha = 1-\gamma / 2$, which is less than one.[11] Since monetary policy is not set solely on the basis of past information, the strength of its impact on output is underestimated by OLS – at least in this simple framework.

Unfortunately, the problem is more serious than simply underestimating the strength of monetary policy. The ordinary least squares approach also tends to overestimate the length of the lags of monetary policy. To illustrate this point, extend the above model to one in which output depends negatively on both the current and first lag of the real interest rate, while the real interest rate responds to information about output in both the current and next periods. In symbols,

$$y_t = -r_t - r_{t-1} + u_t$$

$$r_t = I_t + J_t + v_t$$

where we have introduced the variable J_t which is information available in period t about output in period t+1 and, therefore, correlated with output in period t+1 ($E_t(J_t u_t+1) = \delta$). Estimating the output equation, $y_t = \alpha r_t + \beta r_{t-1} + u_t$, by OLS leads to estimates of α and β which are, on average, smaller in magnitude than their true values, $\alpha = \beta = -1$. If the two errors, u_t and v_t, and the two information variables, I_t and J_t, have zero means and unit variances, the ratios R_α and R_β of the OLS estimates of α and β to their true values are $R_\alpha = 1 - \gamma / 3$ and $R_\beta = 1 - \delta / 3$.[12]

Since information about the current period should be more reliable than information about the future, γ should be greater than δ. Hence, while both coefficients are biased towards zero, the bias is likely to be more serious for α, implying that OLS overestimates the length of the lags of monetary policy.

2.2 How serious is the bias?

We can extend this analysis to examine the bias in the length of the transmission lags from monetary policy to output estimated in Section 1. To derive empirical estimates of the bias, we need to make assumptions about the correlation between the current real interest rate and the current and future unexplained residuals in the non-farm output equation (equation (1)).

We assume a correlation coefficient ρ_k between the current real interest rate, r_t, and the unexplained residual in equation (1) in k-periods time, ε_t+k, of $\rho_k = \gamma^{k+1}$, $0 < \gamma < 1$, $k = 0,\dots,6$. This formula embodies the idea used in the simple models above that the correlation is strongest when the

---

11 The wording in the text is heuristic rather than rigorous. The actual definition of R_α is $R_\alpha = p\lim\ \hat{\alpha}/\alpha$ where $\hat{\alpha}$ is the OLS estimate of α, and $p$lim is the probability limit.

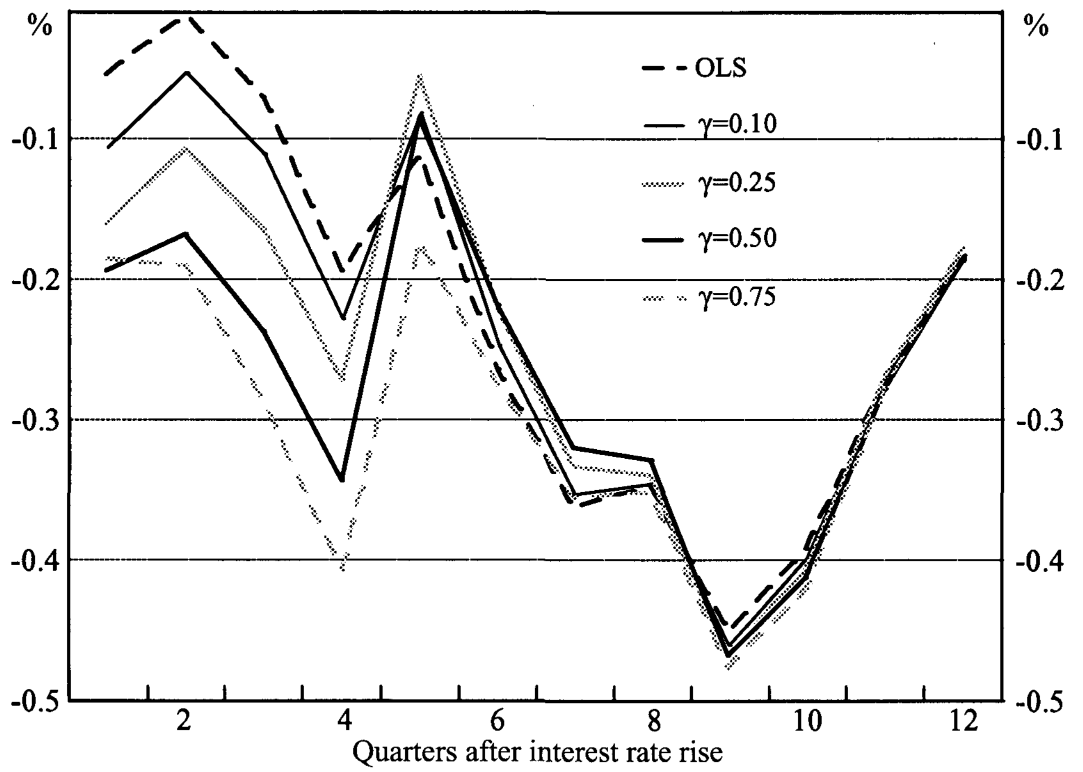
12 Again, the wording in the text is heuristic rather than rigorous. See the previous footnote. The bias we refer to in the text is really the asymptotic bias.

output shock and the real interest rate are contemporaneous and falls as the distance in time between the two variables rises. We derive results for values of γ in the range from zero to 0.75.[13]

Assuming a small value for γ implies that the policy-maker raises real interest rates only slightly when information arrives that current growth will be stronger than is implicit in equation (1), and has virtually no reliable information about future growth other than that predicted by equation (1). By contrast, assuming $\gamma = 0.75$ implies the policy-maker is in possession of good information about both current and future growth shocks, and uses that information to make significant changes to real interest rates.[14]

Figure 4

Impact of monetary policy on four-quarter-ended growth



Note: The figure shows, for the underlying CPI model, the estimated impact on four-quarter-ended growth of non-farm output of a one percentage point rise in the short-term real interest rate adjusting for the bias from a correlation coefficient ρ_k between the current real interest rate and the error in the output equation k-quarters time of: $\rho_k = \gamma^{k+1}$, $k=0,\ldots, 6$, for the values of γ shown.

---

[13] In general, the correlation between the real interest rate and the unexplained residual depends both on the quality of information available about the unexplained residual and the extent to which policy reacts to that information. The assumption of a falling correlation as the distance in time between the two variables rises, could arise either because information is poorer about future residuals, or because policy reacts less to that information.

[14] Assume that the information available to the policy-maker enables him/her to make unbiased forecasts of the current and future shocks to the output equation. Then, for our parameter values, $\gamma = 0.75$ means that when the policy-maker has information implying that the level of output will be half a per cent higher than forecast by the equation in each of the next two years, the short-term real interest rate is immediately raised by two-thirds of one per cent. We judge this to be a significant policy reaction.

Making these assumptions about the correlations between the real interest rate and growth shocks, we can adjust for the bias in the ordinary least squares estimates generated in Section 1. The details of the calculation are presented in Appendix B, and Figure 4 shows results for the underlying CPI model for four values of γ : 0.1, 0.25, 0.5 and 0.75 (both the size and pattern of the bias are very similar in the two models).

As for Figure 3, Figure 4 shows the estimated impact on year-ended non-farm output growth of a one per cent sustained rise in the domestic real interest rate. The figure shows point estimates for the OLS regression summarised in Table 1, and for OLS adjusted using the correlations between growth shocks and real interest rates defined above.

The figure displays both of the features revealed by the simple models described earlier. First, OLS underestimates the average strength of the impact of monetary policy on output. And, second, almost all this underestimation occurs for the impact of monetary policy in the first year, implying that OLS overestimates the average length of the lags of monetary policy.

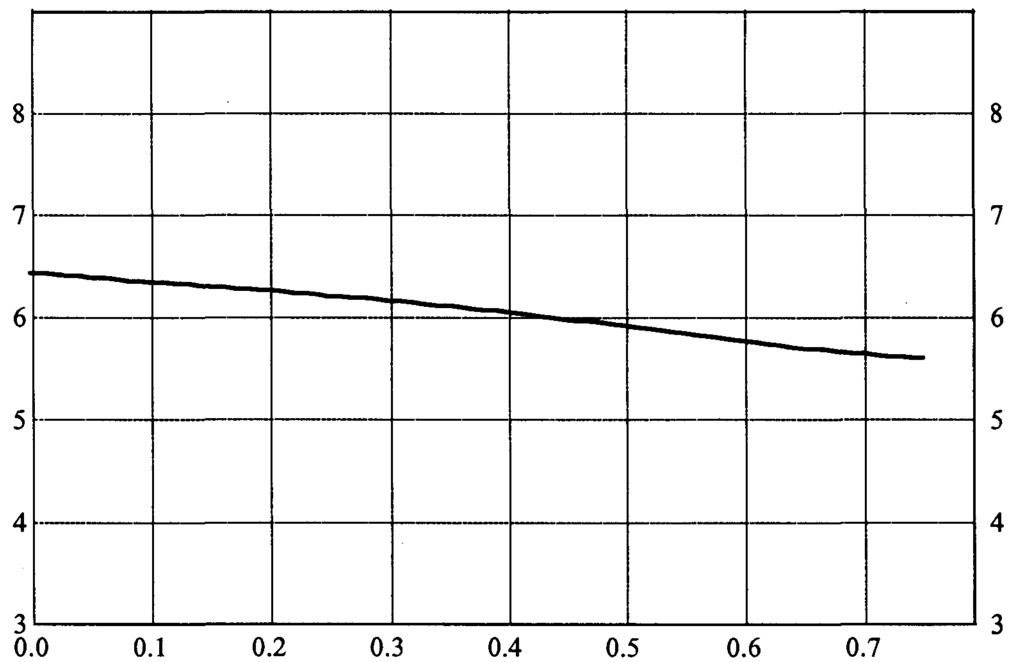
The results in Figure 4 suggest that a moderate amount of correlation between real interest rates and output growth shocks implies that the estimated impact on output growth in the first and second years after a one percentage point rise in the real interest rate is roughly the same, at about one-third of one per cent, falling to about one-sixth of one per cent in the third year. Given the error bands around the original OLS estimates, these numbers are again subject to considerable uncertainty.

We can also calculate the average monetary policy lag as the parameter γ rises from zero to 0.75. Figure 5 shows the results, which confirm that the average lag length falls with rising correlation between the real interest rate and the unexplained residuals in the output equation. It is of interest to note, however, that the effect is not a very powerful one: plausible values of the correlation coefficient imply only small falls in the estimated average monetary policy lag length.

Figure 5

Average lag monetary policy

Adjusting for possible bias in the ordinary least squares regression



Note: The figure shows the average lag in the effect of monetary policy on output growth for the underlying CPI model adjusting for the bias from a correlation coefficient ρ_k between the current real interest rate and the error in the output equation in k-quarters time of $\rho_k=\gamma^{k+1}$, k=0,..., 6, for values ranging from 0 to 0.75.

2.3 Controlling for the policy response using instrumental variables

Rather than hypothesising about the correlation between real interest rates and output growth shocks, an alternative approach is to use instrumental variables to control for the forward-looking nature of monetary policy.

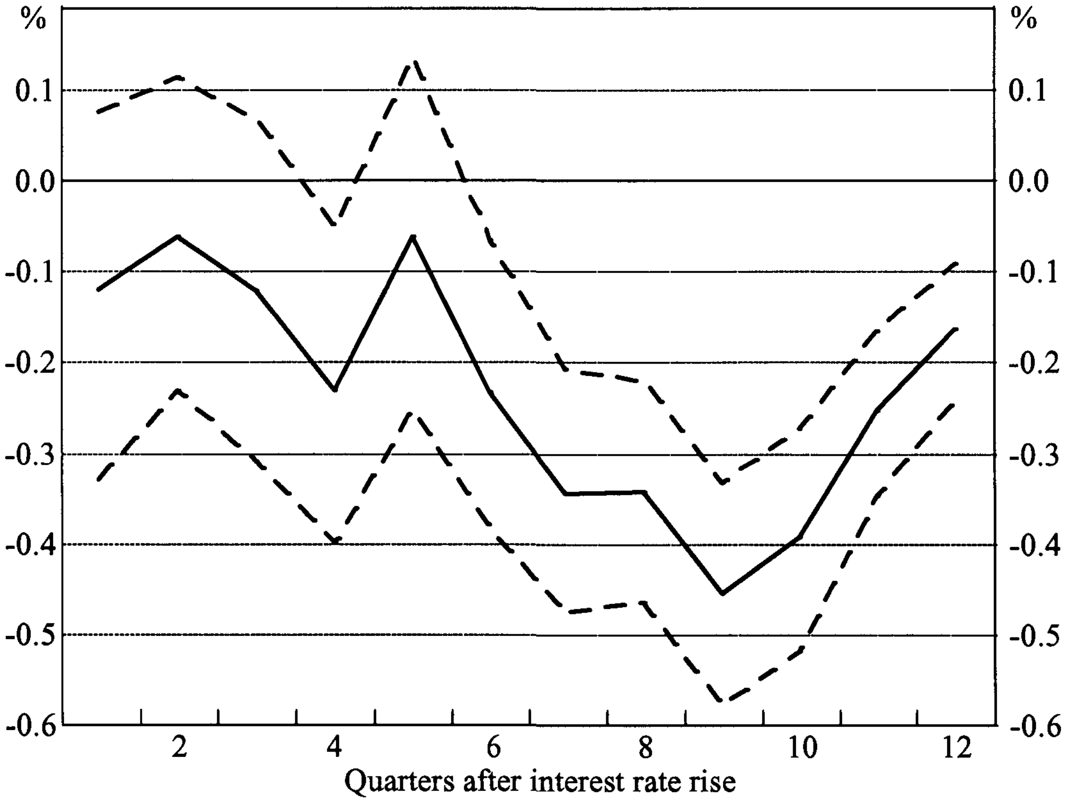
Choosing appropriate instruments for the short-term real interest rate is a difficult task, however, made more so by a change in focus of Australian monetary policy in the 1990s with the introduction of a medium-term inflation target. To allow for this change, we assume that there are two monetary policy regimes, old and new, with the real interest rate responding differently in the two regimes.

Rather than conducting an extensive search for instruments, we confine the list to those variables which should have a strong influence on the short-term real interest rate: inflation, π, and the output gap, *gap*. In the old regime, we also include the current account deficit to GDP ratio, *cad*, as an instrument because we find that this variable has considerable explanatory power.[15] The variables are lagged to reduce the possibility of correlation between them and the error term in the output equation (1).

Figure 6

Impact of monetary policy on four-quarter-ended growth

IV regression simulations



Note: The figure shows, for the underlying CPI model, point estimates and 90% confidence intervals for the impact on four-quarter-ended growth of non-farm output of a one percentage point rise in the short-term real interest rate at the beginning of quarter 1.

---

15 This does not mean that monetary policy "targeted" the current account, as is sometimes suggested, only that policy responded to changes in demand conditions which were reflected in changes in the current account deficit.

For brevity, we again show only the results using underlying inflation (although the results are similar for headline inflation). We use the fitted values from the regression,

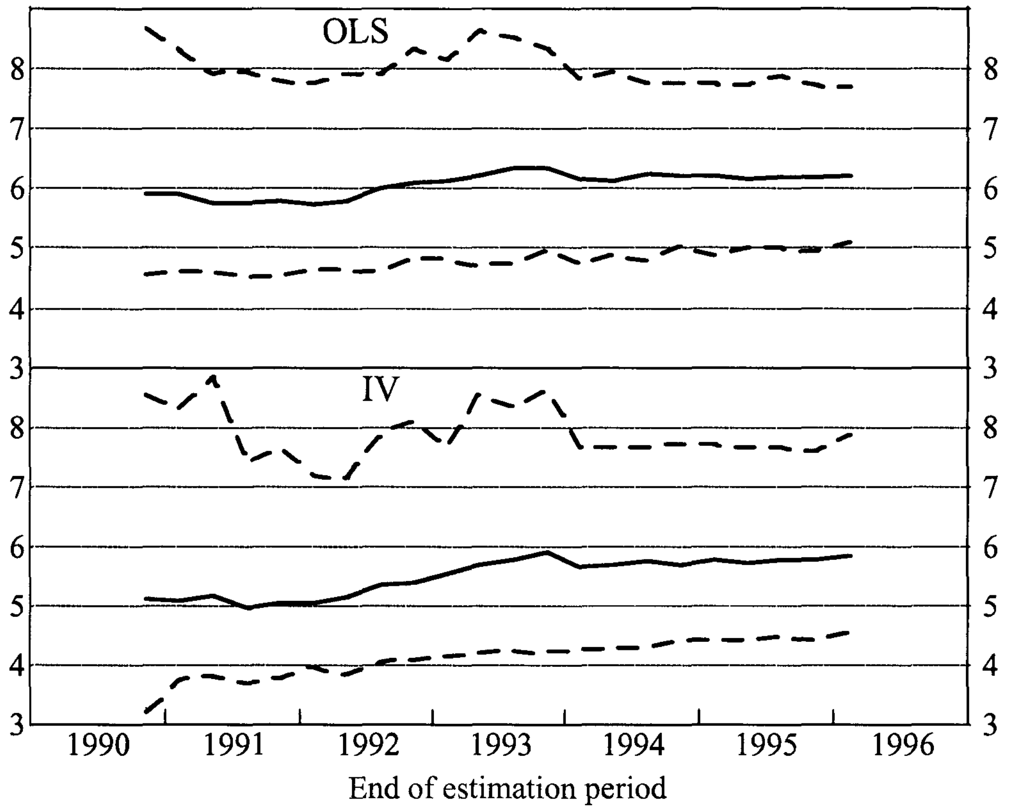
$$r_t = a^0 d^0_{t-1} + b^0 \pi^0_{t-1} + c^0 gap^0_{t-1} + e^0 cad^0_{t-1} + a^n d^n_{t-1} + b^n \pi^n_{t-1} + c^n gap^n_{t-1} + u_t \quad (3)$$

to instrument for the contemporaneous real interest rate in the non-farm output equation (1). The notation implies, for example, that π^0_{t-1} is the (four-quarter-ended) underlying inflation rate in period t-1 when t is in old regime and 0 when t is in the new regime. d^0_{t-1} and d^n_{t-1} are dummy variables for the two regimes and *gap* is defined as the deviation of non-farm output from a linear trend. The start of the new regime is the first quarter of 1990, chosen to maximise the explanatory power of the equation. The instruments together explain 78% of the variance of the short-term real interest rate over the estimation period, 1980Q3 to 1996Q1.[16]

Figure 7

Average lag of monetary policy

Extending the estimation period through the 1990s



Note: The figure shows recursive point estimates and 90% confidence intervals for the average lag in the effect of monetary policy on output growth for the underlying CPI model, using both OLS and IV estimation techniques.

From this instrumental variable (IV) regression, we derive the impact of a sustained one per cent rise in the short-term real interest rate on the year-ended growth of non-farm output. The

[16] Since policy-makers set the short-term real interest rate partly on the basis of information about current output not available to the econometrician estimating equation (1), we should expect positive correlation between the errors in equations (1) and (3). The estimated correlation coefficient between the errors is, at 0.08, positive, but small.

results are shown in Figure 6[17] and are not markedly different from the ordinary least squares results using the underlying CPI model in Figure 3. The estimated contractionary impact in the first year is somewhat stronger in the IV regression than in the OLS one, but it is also less precisely estimated. In common with the OLS results, the estimated impact on output growth in the second year is somewhat stronger than in the first year, but the difference is again statistically insignificant.[18]

2.4 Have the lags of monetary policy changed over time?

To examine whether the lags of monetary policy have changed over time, we conduct recursive estimation on the underlying CPI model, using both ordinary least squares and instrumental variables regressions. The recursive estimation fixes the starting date at the beginning of the full sample, 1980Q3, and extends the end of the estimation period from 1990Q3 to 1996Q1, one quarter at a time. Figure 7 shows how the average monetary policy lag (defined by equation (2)) changes as the estimation period is lengthened through the 1990s. As before, both point estimates and 90% confidence intervals are shown.

Table 2
Model comparison of the effect on output growth of a one percentage point rise in short-term interest rates
In percentage points[1]

Model	First year	Second year	Third year
Australia			
Current model[2]	- 0.35	- 0.33	- 0.18
Murphy[3]	- 0.51	- 0.31	+ 0.08
TRYM[3]	- 0.37	- 0.37	- 0.07
United States[4]			
MPS	- 0.20	- 0.70	- 1.10
DRI	- 0.47	- 0.53	- 0.13
FAIR	- 0.24	- 0.25	+ 0.03
FRBSF	- 0.55	- 0.19	+ 0.04
VAR	- 0.64	- 0.26	+ 0.08

[1] For the Australian models, the policy experiment involves raising the short-term real interest rate, while for the US models, the short-term nominal rate is raised. Given the sluggish adjustment of inflation, this difference should not matter much. [2] The results for the current model are those for the underlying CPI model, bias-adjusted assuming $\gamma = 0.5$. [3] See Murphy (1995) and The Treasury (Australia) (1996) for a description of the models. Assuming the models are linear for small changes, the results are derived by adding up the impact on output of a series of nominal shocks which together imply a one percentage point rise in the short-term real interest rate for 12 quarters. The results using the TRYM model should in no way be regarded as being Treasury analyses of the effect of a given policy change or as having the sanction of the Treasury, the Treasurer or the Commonwealth Government. [4] The MPS model is maintained by the staff of the Federal Reserve Board, the DRI model is commercially available, the FAIR model is maintained by Ray Fair of Yale University, the FRBSF model is maintained by the Federal Reserve Bank of San Francisco, while the VAR model is a standard vector autoregressive model (see Rudebusch (1995) for further details).

---

[17] As before, the confidence intervals are estimated using a Monte Carlo procedure described in Appendix A.

[18] It would be of interest to instrument not only for the contemporaneous real interest rate but also for some of its lags. Unfortunately, given the strong correlation between the real interest rate and its lags and the fact that the instruments are not very good, this exercise did not yield any useful information.

The results do not suggest any shortening of the lags of monetary policy over the 1990s. As the estimation period is lengthened, there is a slight rise in the average monetary policy lag length, which is somewhat larger in the IV than in the OLS regressions. Needless to say, these results are not statistically significant because the confidence intervals are so wide.

2.5 Comparison with other estimates of the lags of monetary policy

Table 2 shows the effect on output growth in the first, second and third years after a rise in the short-term interest rate for a range of models estimated for the Australian and US economies.

There is qualitative agreement between the models that output growth falls in both the first and second years after a rise in the short-term interest rate, although there is no such agreement about the third year. Even for the first two years, however, the quantitative estimates differ quite substantially between models.[19] This is simply another manifestation of the technical difficulties associated with estimating the lags of monetary policy.

Conclusions

We draw three broad conclusions about the transmission lags from the short-term real interest rate to output in the Australian economy.

First, there is strong econometric evidence that the level of the short-term real interest rate has a sizeable, and statistically significant, impact on output in the Australian economy. Ordinary least squares estimation suggests that a one percentage point rise in the short-term real interest rate lowers output growth by one-fifth to one-quarter per cent in the first year, one-third per cent in the second year and one-sixth per cent in the third year, although these estimates are subject to considerable uncertainty.

Second, we examine the implications of monetary policy being based not only on past information, but also on current and expected future developments in the economy. We show that, as a consequence, the strength of monetary policy's impact on output is underestimated by ordinary least squares estimation, while the length of the monetary policy lags are overestimated. Under plausible assumptions, the underestimation is concentrated in the first year, while the estimated impact on output growth in the second and third years is largely unaffected. This leads us to conclude that the effect of a one percentage point rise in the short-term real interest rate on output growth is probably similar in the first and second years – at about one-third per cent in each year – declining to about one-sixth per cent in the third year, with an average lag of about five or six quarters. All these estimates are, however, subject to considerable uncertainty.

Thirdly, we look for evidence of changes in the lags of monetary policy over the 1990s. Our results, based on both ordinary least squares and instrumental variable estimation, do not suggest any shortening in these lags. If anything, they suggest that the lags of monetary policy may have become slightly longer over time, although our estimates are not sufficiently precise to be very confident of this conclusion.

---

[19] In the case of the Murphy model, the stronger contractionary impact in the first year may be partly a consequence of the assumption of uncovered interest parity which implies an initial jump appreciation of the real exchange rate and, therefore, a significant contractionary impact on output via net exports.

Appendix A: Monte Carlo procedure

This appendix outlines the Monte Carlo procedure used to generate confidence intervals for the OLS, IV and recursive regressions.

A.1 Ordinary least squares regressions

The non-farm output equation, rewritten here for convenience, is:

$$\Delta y_t = \alpha + \sum_{j=0}^{6}\beta_j r_j + \left[\gamma_2 \Delta f_{t-2} + \gamma_4 \Delta f_{t-4}\right] + \delta y_{t-1} + \chi w_{t-1} + \sum_{j=0}^{1}\phi_j \Delta w_{t-j} + \varepsilon_t \tag{A1}$$

which may be simplified to:

$$\Delta y_t = \delta y_{t-1} + N_t \lambda + \varepsilon_t \tag{A2}$$

where N_t is the vector of explanatory variables excluding y_{t-1}.

A sustained 1% rise in the real interest rate leads to an effect on the level of output after j quarters (m_j) of:

$$m_0 = 0$$

$$m_j = (1+\delta)m_{j-1} + \sum_{i=0}^{k}\beta_i \text{ where } k = \min(j - 1, 6) \tag{A3}$$

$$m_\infty = -\sum_{i=0}^{6}\beta_i / \delta.$$

Estimating equation (A2) by OLS over the 63 quarters 1980Q3 to 1996Q1 leads to parameter estimates $\hat{\delta}$ and $\hat{\lambda}$, and an estimate of the standard deviation of the errors, $\sigma_\varepsilon = 0.56$, for both the underlying and headline models. The Monte Carlo distribution is then generated by running 1,000 trials with each trial, i, proceeding as follows:

1. draw a sequence of observations $\left\{e_t^i\right\}_{t=1}^{63}$ from a normal distribution with mean 0 and variance σ_ε^2;
2. generate sequences of synthetic data $\left\{\Delta y_t^i\right\}_{t=1}^{63}, \left\{y_t^i\right\}_{t=1}^{63}$ using $\Delta y_t^i = \hat{\delta} y_{t-1}^i + N_t \hat{\lambda} + \varepsilon_t^i$ and $y_t^i = y_{t-1}^i + \Delta y_t^i$, where $\hat{\delta}$ and $\hat{\lambda}$ are from the OLS estimation using the original data;
3. use the synthetic data to estimate the equation $\Delta y_t^i = \delta^i y_{t-1}^i + N_t \lambda^i$, by OLS and hence generate parameter estimates $\hat{\delta}^i$ and $\hat{\lambda}^i$; and
4. with these parameter estimates, use equation (A3) to calculate, for this i^{th} iteration, the effect of a 1% rise in the real interest rate on the level of output (m_j^i, j = 1,...,12,..., ∞) and the year-ended growth rate of output ($m_j^i - m_{j-4}^i$) after j quarters.

The figures in the text show the 5th, 50th and 95th percentile values for the effect on the level of output, m_j^i, and on the year-ended growth rates, $m_j^i - m_{j-4}^i$.

A.2 Instrumental variable regressions

The policy reaction function, rewritten for convenience, is:

$$r_t = a^0 d_{t-1}^0 + b^0 \pi_{t-1}^0 + c^0 gap_{t-1}^0 + e^0 cad_{t-1}^0 + a^n d_{t-1}^n + b^n \pi_{t-1}^n + c^n gap_{t-1}^n + u_t \quad \text{(A4)}$$

Estimating the underlying CPI version of equation (A4) by OLS over the 63 quarters 1980Q3 to 1996Q1 leads to fitted values $\hat{r}_t$, and an estimate of the standard deviation of the errors, $\sigma_u = 1.32$. Diagnostic tests on the sample errors reveal strong signs of first-order autocorrelation, with an estimated autocorrelation coefficient, $\hat{\rho} = 0.31$.

Estimating equation (A2) by IV, using $\hat{r}_t$ as an instrument for r_t over the period 1980Q3 to 1996Q1 leads to parameter estimates $\tilde{\delta}$ and $\tilde{\lambda}$, and an estimate of the variance-covariance matrix of the errors from equations (A2) and (A4), $\tilde{V}$. The Monte Carlo distribution is then generated by running 1,000 trials with each trial, i, proceeding as follows:

1. draw two sequences of observations $\{\varepsilon_t^i\}_{t=1}^{63}$ and $\{u_t^i\}_{t=1}^{63}$ from a bivariate normal distribution with mean 0 and covariance matrix $\tilde{V}$, such that $u_t^i = \hat{\rho} u_{t-1}^i + \eta_t^i$, where η_t^i are independent and identically distributed;

2. generate sequences of synthetic data $\{\Delta y_t^i\}_{t=1}^{63}$, $\{y_t^i\}_{t=1}^{63}$ using $\Delta y_t^i = \tilde{\delta} y_{t-1}^i + N_t \tilde{\lambda} + \varepsilon_t^i$ and $y_t^i = y_{t-1}^i + \Delta y_t^i$, where $\tilde{\delta}$ and $\tilde{\lambda}$ are from the IV estimation using the original data;

3. generate a sequence of synthetic data $\{r_t^i\}_{t=1}^{63}$ according to $r_t^i = \hat{r}_t + u_t^i$. Re-estimate equation (A4) by OLS using r_t^i instead of r_t and obtain a new set of fitted values, $\breve{r}_t^i$;

4. estimate the equation $\Delta y_t^i = \delta^i y_{t-1}^i + N_t \lambda^i$ by IV, using $\breve{r}_t^i$ as an instrument for r_t, and hence generate parameter estimates $\breve{\delta}^i$ and $\breve{\lambda}^i$; and

5. with the parameter estimates $\breve{\delta}^i$ and $\breve{\lambda}^i$, use equation (A3) to calculate, for this i^{th} iteration, the effect of a 1% rise in the real interest rate on the year-ended growth rate of output, $m_j^i - m_{j-4}^i$, after j quarters.

The figures in the text show the 5th, 50th and 95th percentile values for the year-ended growth rates.

A.3 Recursive regressions

For the recursive regressions, a new Monte Carlo distribution is estimated from 1,000 trials after each new quarter of data is added.

Appendix B: Estimating the bias from ordinary least squares

It is convenient to rewrite the model for non-farm GDP growth, equation (1) in the text, as:

$$\Delta y_t = \sum_{j=0}^{6} \beta_j r_{t-j} + \delta Z_{t-1} + W_t \phi + \varepsilon_t, \qquad \text{(B1)}$$

where $Z_{t-1} = y_{t-1} - \chi^* w_{t-1}$ (χ^* is the cointegrating vector between y and w) and W_t is the matrix of exogenous variables, $W_t = \left[1 \Delta w_t \;\; \Delta w_{t-1} \;\; \Delta f_{t-2} \;\; \Delta f_{t-4}\right]$. Equation (B1) may be further simplified to:

$$\Delta y_t = X_t \alpha + W_t \phi + \varepsilon_t, \qquad \text{(B2)}$$

where $X_t = \left[r_t \;\; r_{t-1} \;\; r_{t-2} \;\; r_{t-3} \;\; r_{t-4} \;\; r_{t-5} \;\; r_{t-6} \;\; Z_{t-1}\right]$ is the matrix of regressors presumed to be correlated with the disturbance term, ε_t.

OLS on equation (B2) yields the following estimate for α,

$$\hat{\alpha}_{OLS} = \alpha + \left(X' M_W X\right)^{-1} X' M_W \varepsilon \qquad \text{(B3)}$$

where $M_W = I - W\left(W'W\right)^{-1} W'$.

Now,

$$\begin{aligned} p\lim \frac{X' M_W \varepsilon}{T} &= p\lim \frac{X'\varepsilon}{T} - p\lim X'W\left(W'W\right)^{-1} \frac{W'\varepsilon}{T} \\ &= p\lim \frac{X'\varepsilon}{T} - p\lim X'W\left(W'W\right)^{-1} p\lim \frac{W'\varepsilon}{T} \\ &= p\lim \frac{X'\varepsilon}{T} \end{aligned}$$

as, with W exogenous, $p\lim \frac{X'\varepsilon}{T} = 0$. In the limit, the true value of the vector α is then:

$$\alpha = \hat{\alpha}_{OLS} - p\lim \left(\frac{X' M_W X}{T}\right)^{-1} p\lim \frac{X'\varepsilon}{T} \qquad \text{(B4)}$$

We presume that the short-term real interest rate, r_t, can be expressed as:

$$r_t = f\,(\text{exogenous variables}) + u_t \qquad \text{(B5)}$$

where "exogenous" implies uncorrelated with the error term in equation (B2) and u_t is determined by the policy-maker on the basis of information about current and future output not available to the econometrician estimating (B2).

As explained in the text, the correlation coefficient between real interest rates and the error term in equation (B2) is assumed to be a geometrically declining function of the lag of the real interest rate, with no correlation after the sixth lag. The covariance between Z_{t-1} and ε_t (which is identical to the covariance between y_{t-1} and ε_t) is denoted $\sigma_\varepsilon \sigma_u \theta$ and is derived below. In symbols we have:

$$p\lim \frac{X'\varepsilon}{T} = \sigma_\varepsilon \sigma_u \left[\gamma\,\gamma^2\,\gamma^3 \ldots \gamma^7\,\theta\right]', \tag{B6}$$

where σ_ε and σ_u are estimates of the standard deviations of the errors in equations (B1) and (B5). For the underlying model, $\sigma_\varepsilon = 0.56$ while, for σ_u, we use the value derived from estimating equation (3) in the text (which is a simple version of equation (B5)). This gives the estimate $\sigma_u = 1.32$.

Now define the variables C_i, $i = 0,\ldots,6$ by:

$$C_i = \sigma_\varepsilon \sigma_u \sum_{j=0}^{6-i} \beta_j \gamma^{i+j+1} \tag{B7}$$

Denote the covariance between ε_t and y_{t-i} as PL_i, and between ε_t and Δy_{t-i} as PC_i. The model, equation (B2), implies the recursive structure,

$$\begin{aligned} PL_7 &= 0 \\ PC_6 &= C_6 \\ PL_i &= PL_{i+1} + PC_i \\ PC_i &= C_i + \delta PL_{i+1} \end{aligned} \tag{B8}$$

We require $\theta = PL_1 / \sigma_\varepsilon \sigma_u$, which is a function of the true vector α. For given γ in the range 0 to 0.75, we proceed as follows. First, we use the sample value of $\left(\frac{X'M_W X}{T}\right)^{-1}$ as our estimate of $p\lim\left(\frac{X'M_W X}{T}\right)^{-1}$. (This requires an estimate of the cointegrating vector, χ^*, between y and w; we use the OLS estimate for this.) Next, we use the OLS estimate, $\hat{\alpha}_{OLS}$, to generate an estimate $\hat{\theta}$ via equation (B8). We now have an estimate for $p\lim \frac{X'\varepsilon}{T}$ via equation (B6). This enables us to generate an estimate, $\hat{\alpha}$, of the "true" vector α via equation (B4). We now iterate: $\hat{\alpha}$ implies a new estimate for θ, $\hat{\hat{\theta}}$, which, in turn, implies a new estimate for α, $\hat{\hat{\alpha}}$. This process is continued until it converges, yielding $\tilde{\alpha}$. The estimated response to a permanent 1% increase in the real interest rate on year-ended growth shown in Figure 4 and on the average lag of monetary policy shown in Figure 5 are generated using $\tilde{\alpha}$.

References

Debelle, G. and B. Preston (1995): "Consumption, Investment and International Linkages". Reserve Bank of Australia, *Research Discussion Paper*, No. 9512.

de Brouwer, G. and J. Romalis (1996): "External Influences on Output: An Industry Analysis". Reserve Bank of Australia, *Research Discussion Paper*, No. 9612.

de Roos, N. and B. Russell (1996): "Towards an Understanding of Australia's Co-movement with Foreign Business Cycles". Reserve Bank of Australia, *Research Discussion Paper*, No. 9607.

Dixit, A. and R. Pindyck (1994): *Investment Under Uncertainty*, Princeton University Press, Princeton.

Eichenbaum, M. and C.L. Evans (1995): "Some Empirical Evidence on the Effects of Shocks to Monetary Policy on Exchange Rates". *Quarterly Journal of Economics*, 110(4) pp. 975–1,009.

Grenville, S. (1996): "Recent Developments in Monetary Policy: Australia and Abroad". *Australian Economic Review*, 113(1), pp. 29-39.

Gruen, D. and G. Shuetrim (1994): "Internationalisation and the Macroeconomy", in P. Lowe and J. Dwyer (eds), *International Integration of the Australian Economy*. Proceedings of a Conference, Reserve Bank of Australia, Sydney, pp. 309-363.

Jorgenson, D. and J. Stephenson (1967): "Investment Behaviour in U.S. Manufacturing 1947-60". *Econometrica*, 35(2), pp. 169-220.

Kortian, T. and J. O'Regan (1996): "Australian Financial Market Volatility: An Exploration of Cross-Country and Cross-Market Linkages". Reserve Bank of Australia, *Research Discussion Paper*, No. 9609.

Lowe, P. (1995): "The Link Between the Cash Rate and Market Interest Rates". Reserve Bank of Australia, *Research Discussion Paper*, No. 9504.

McTaggart, D. and T. Hall (1993): "Unemployment: Macroeconomic Causes and Solutions? or Are Inflation and the Current Account Constraints on Growth?" *Bond University Discussion Paper*, No. 39.

Murphy, C. (1995): "Murphy Model 2". *1995 Symposium on Modelling and Control of National and Regional Economies*, The Institute of Engineers, Australia, June.

Pesaran, M. H., Y. Shin and R. J. Smith (1996): "Testing for the Existence of a Long-Run Relationship", *mimeo*.

Rudebusch, G. (1995): "What are the Lags in Monetary Policy?" *Federal Reserve Bank of San Francisco Weekly Newsletter*, February 3, Federal Reserve Bank of San Francisco.

Shapiro, M. (1986): "The Dynamic Demand for Capital and Labor". *The Quarterly Journal of Economics*, 101(3), pp. 513-542.

The Treasury (Australia) (1996): *TRYM Model: Treasury Macroeconomic Model of the Australian Economy; TSP version*, ABS Cat. No. 1364.0.15.001, ABS / The Treasury, Canberra.

Comments on: "The lags of monetary policy" by David Gruen, John Romalis and Naveen Chandra

by Benjamin H. Cohen

This paper addresses an important and frequently encountered question: how do we measure the effects of monetary policy moves on output, when those moves themselves reflect contemporaneous developments in output? The reflexive response by many macroeconomists would probably be to construct a vector auto-regression (VAR), which gets around the simultaneity issue by using the information in *lagged* output and policy variables (and usually other variables as well). This works well enough if the time-series structures of the variables are sufficiently stable, but it ignores the potential information that one could get from *contemporaneous* movements. Consider the attention that market observers and makers of monetary policy devote to incomplete, but very timely, market indicators – such as retail sales figures, surveys of business and consumer confidence, or even anecdotal reports – at the expense of more accurate figures that come out with longer delays. Conversely, there are certainly some elements of economic activity, such as housing investment, which respond to short-term interest rate movements at lags of less than one quarter.

Gruen, Romalis, and Chandra begin by estimating a simple OLS model for Australian GDP growth as a function of lagged real rates, farm output growth, and lagged Australian and US output. Depending on the inflation variable used, the models have adjusted R-squared's of 0.60 or 0.59. The models imply an "average lag length" (a kind of duration concept, in which the number of quarters of declining output is weighted by the size of the decline) of 6.4 or 5.8 quarters.

The authors then explore two methods for including contemporaneous mutual output and real-rate effects. First, they make various assumptions about policymakers' response to unexpected movements in current output. A stronger policy response usually leads to a larger GDP fall in the first year, but has little effect after that, and the average lag of monetary policy effects is surprisingly resilient to a range of estimated values for this response. Second, they address the simultaneity problem using six instruments for the real short-term interest rate: lagged inflation, the lagged output gap, the lagged current account deficit, a dummy variable indicating a change in Australia's monetary policy regime towards an inflation target, and two interaction terms reflecting the increased emphasis on inflation and the output gap under the inflation-targeting regime. As it happens, their results for the effect of a sustained one-percentage-point rise in the short rate on output are almost the same using this technique as using OLS. Perhaps this suggests that the simultaneity problem, which (as the authors show) ought to have biased the OLS coefficients considerably downwards, is not as bad as was thought.

The authors find that a one-percentage-point real rate rise reduces growth by about a third of a percentage point in the first two years, and about a fifth of a percentage point in the third year. They show that, for the first two years, this is roughly comparable to the results of other models for Australia and close to the mid-point of the wider range of estimates for the US economy, though they find a somewhat deeper recessionary impact in the third year than most of the other models.

Two further points:

- In their benchmark OLS output equation, the authors use the lagged levels of Australian and US log GDP as error-correction terms. They state in a footnote that the US term does not play the role of a trend, because a trend term was insignificant when added to the regressions. I would be more convinced if the level terms had simply been detrended to start with; this would seem to better incorporate the business-cycle effects that the authors want to capture.
- The authors show that the estimated lags in monetary policy effects lengthen as the end of the estimation period is extended from 1990 to 1996. This suggests that the "centre of gravity" of the output decline has moved outward somewhat. It would be interesting, however, to know

whether the overall decline – the area between the actual-output and trend-output curves – has fallen. In other words, what has happened to the overall effectiveness of monetary policy? This seems a relevant issue in current circumstances, when significant loosenings of monetary policy in Japan and a number of European countries seem, after a lag, to have produced genuine but disappointingly weak recoveries.

Effects of a single European monetary policy: simulations with the multi-country model of the Deutsche Bundesbank

Wilfried Jahnke and Bettina Landau

Introduction

In the spring of 1998 it will be decided whether the European Monetary Union (EMU) will start in 1999 and which countries will participate. In the meantime it is completely open which countries will qualify then. One of the pre-conditions set in the Maastricht Treaty is at least a two-year participation in the exchange rate mechanism (ERM) of the European Monetary System without severe exchange rate tensions. For that reason Italy has returned to the ERM in November 1996 to reserve the possibility of joining the EMU from the beginning. Principally all countries of the European Union now participating in the exchange rate mechanism of the European Monetary System have a chance to fulfil the convergence criteria laid down in the Maastricht Treaty. It may even be possible that the United Kingdom and Denmark which have been granted an opting-out clause in the Maastricht Treaty will participate in the European Monetary Union. But the necessary regime shift from national monetary policies to a common monetary policy of the European Central Bank cannot be realised overnight. On the way to the European Monetary Union it seems necessary to coordinate national monetary policies converging to the likely single European monetary policy because otherwise the changes in monetary policy instruments will become too large and too abrupt.[1] Coordination of monetary policies will be the more necessary when the decisions on the participating countries have been taken in 1998.

The common monetary policy will result in a single short-term interest rate within the European Monetary Union, irrespective of whether the European Central Bank follows a strategy of monetary or of inflation targeting.[2] At present (January 1997), an assessment of macro-economic effects of the transition to a single monetary policy in EMU is not possible without some rather heroic and very strong assumptions. In the following analysis with the macro-econometric multi-country model of the Deutsche Bundesbank it has been assumed, therefore, that:

- EMU starts at the beginning of 1999,
- all six European countries contained in the model, i.e. Belgium, France, Germany, Italy, the Netherlands and the United Kingdom, will fulfil the convergence criteria and participate in EMU from the beginning,[3]
- the conversion rates to the Euro will be identical to the present exchange rates (which are not far from ERM central rates for the participating currencies).[4] There will be no realignments within the ERM in the years 1997 and 1998.

---

1 Actually, the strengthening of the coordination of Member States' national monetary policies is a main task of the European Monetary Institute in Stage II of the Maastricht process.

2 The necessary strategy decisions and the operational framework for the single monetary policy have been described in EMI (1997).

3 However, the reluctance of the British Government to rejoin the ERM and to make the Bank of England fully independent may indicate that the UK will not participate in EMU in the first round.

4 On other possible procedures to fix the conversion rates see DeGrauwe (1996).

This means that a fully credible fixed exchange rate system is already in effect in advance of EMU. This credibility may be strengthened further by an early announcement of the conversion rates and a commitment to consistent underlying monetary and economic policies.[5] But the still existing differences in short-term interest rates point to lasting exchange rate risks, especially in the case of the Italian lira and the pound sterling. At the moment, risk premiums are very low between the currencies of the so-called "core"-countries Belgium, France, Germany and the Netherlands. These countries already form a quasi monetary union approximating an optimal currency area. But risk premiums are distinctly higher against the Italian lira and the pound sterling.[6]

The paper is organised as follows. Following a short description of the model the effects of a transition from national monetary policies to a single European monetary policy have been simulated. This common policy has been formulated in two alternatives: first as an adjustment to German short-term interest rates and second as an adjustment to an average European interest rate level. The paper closes with some tentative conclusions.

1. Overview of the model

In analysing the transition from the European Monetary System to the European Monetary Union the macro-econometric multi-country model of the Deutsche Bundesbank is used. This model consists of nine country models for the United States, Japan, Germany, the United Kingdom, France, Italy, Canada, the Netherlands, Belgium and a foreign trade sector. The German country model has nearly 150 equations, whereas the other country models have between 53 and 56 equations. This adds up to a total of nearly 640 equations (see Table 1). The German model, which has been estimated for Germany as a whole, is much more disaggregated than the others and therefore contains more structural details of the economy. All of the country models are built on long-run neo-classical theory with short-run rigidities in the labour and goods markets and an adaptive expectation formation. But only some of the behavioural equations in a country model are decisive. These are especially the equations for real private consumption, real private investment, real imports, labour demand, potential output, effective wages, price deflators, real money demand, long-term interest rates and exchange rates (Table 2 gives a simplified version of a country model). Thus the country models belong more to the small size scale. Most of the other equations are included to facilitate the description and the simulation of the model. In many cases the dynamics of the equations are based on an error correction mechanism.[7]

The exchange rate of ERM currencies against the Deutsche mark is modelled as a quasi-fixed exchange rate. This implies that the short-term interest rate in the respective countries equals the German short-term interest rate plus a foreign exchange risk premium. The exchange rates of the mark and of non-ERM currencies against the US dollar are explained by equations based on interest rate differentials against the United States and purchasing power parity.[8] The short-term interest rates in these countries are either described by an autoregressive equation or can be used as the exogenous monetary policy instrument or can be specified in a monetary policy reaction function.

---

5 See Bayoumi (1996).

6 Other European countries like Austria, Denmark, Spain and so forth as well as their currencies are not contained in the model used.

7 Deutsche Bundesbank (1996).

8 Jahnke (1996).

Table 1
Size and structure of the multi-country model: number of equations

Sector	US	JP	DE	UK	FR	IT	CA	NL	BE	FT	Total
I. Aggregate demand											
Estimated equations	4	4	7	4	4	4	4	4	4		39
Identities	11	11	20	11	11	11	11	11	11	31	139
Total	15	15	27	15	15	15	15	15	15	31	178
II. Aggregate supply											
Estimated equations	10	10	14	9	8	9	10	10	9		89
Identities	6	6	19	7	8	7	6	6	7		72
Total	16	16	33	16	16	16	16	16	16		161
III. Factor cost and price deflators											
Estimated equations	6	6	14	6	6	6	6	6	6		62
Identities	5	7	7	5	5	7	5	5	5	17	68
Total	11	13	21	11	11	13	11	11	11	17	130
IV. Government sector											
Estimated equations	3	3	9	3	3	3	3	3	3		33
Identities	4	4	10	4	4	4	4	4	4		42
Total	7	7	19	7	7	7	7	7	7		75
V. Money, interest and exchange rates											
Estimated equations	4	5	31	4	5	5	5	5	5		69
Identities	1		18		2		1	2	2		26
Total	5	5	49	4	7	5	6	7	7		95
Total model											
Estimated equations	27	28	75	26	26	27	28	28	27		292
Identities	27	28	74	27	30	29	27	28	29	48	347
Total	54	56	149	53	56	56	55	56	56	48	639
Exogenous variables	2	2	36	2	4	4	4	5	3	4	66

Note: BE: Belgium, CA: Canada, DE: Germany, FR: France, FT: Foreign trade, IT: Italy, JP: Japan, NL: Netherlands, UK: United Kingdom, US: United States.

Table 2
Simplified version of a country model

I. Aggregate demand		
(1)	Private consumption:	$\frac{C}{WO} = C\left(\frac{YV}{WO * p}, r - \dot{p}^{e}\right)$
(2)	Labour supply:	$\frac{E}{WO} = E\left(\frac{WA}{WO}, \frac{w(1-t_d)}{p}\right)$
(3)	Real final demand:	$Y = C + I + \frac{G}{P} + X$
(4)	National income:	$VE = Y * p - d * K * p - TI - IM * m$
(5)	Disposable income:	$YV = VE + TR - TD$
II. Aggregate supply		
(6)	Optimal real capital stock:	$K* = K*\left(Y, \frac{p*(1-t_i)}{c}\right)$
(7)	Private investment:	$I = I(K, K*)$
(8)	Labour demand:	$A = A\left(Y, \frac{p*(1-t_i)}{w}\right)$
(9)	Imports:	$IM = IM\left(Y, \frac{p*(1-t_i)}{m}\right)$
(10)	Potential output:	$Y* = Y*\left(E, \frac{E-A}{E}, K\right)$
(11)	Real capital stock:	$K = (1-d)*K_{-1} + I$
III. Factor costs and price deflators		
(12)	Wage rate:	$\dot{w} = w\left(\dot{p}^{e}, \frac{E-A}{E}\right)$
(13)	Price deflator of domestic demand:	$\dot{p} = p\left(\dot{w}, \dot{m}, \frac{Y-IM}{Y^{*}}, \frac{p^{*}}{p}\right)$
(14)	Expected inflation rate:	$\dot{p}^{e} = \alpha\dot{p}^{e}_{-1} + (1-\alpha)\dot{p}_{-1}$
(15)	User costs of capital:	$c = p*\left[r*(1-t_d) - \dot{p}^{e} + d\right]$

Table 2
Simplified version of a country model (continued)

IV. Government sector

(16)	Direct taxes:	$TD = t_d * VE$
(17)	Indirect taxes:	$TI = t_i * Y * p$

V. Money, interest rates and exchange rates

(18)	Money stock:	$\frac{M}{p} = M(Y - IM, r)$
(19)	Long-term price level:	$p^* = p^*\left(M, Y^*\right)$
(20)	Long-term interest rate:	$r = r\left(\dot{p}^e\right)$
(21)	Exchange rate:	$e = e\left(rsf - rs, \frac{p}{pf}\right)$

Variables

A	Labour demand (employment)
C	Real private consumption
c	User costs of capital
d	Depreciation rate
E	Labour supply (labour force)
e	Exchange rate against US dollar
G	Nominal government expenditure
I	Real private investment
IM	Real imports
K	Real capital stock
K^*	Optimal real capital stock
M	Money stock
m	Price deflator of imports
p	Price deflator of domestic demand
$\dot{p}^e$	Expected inflation rate
p^*	Long-term price level
pf	Foreign price deflator of domestic demand

Table 2
Simplified version of a country model (continued)

r	Long-term interest rate
rs	Short-term interest rate
rsf	Foreign short-term interest rate
TD	Direct taxes
t_d	Direct tax rate
TI	Indirect taxes
t_i	Indirect tax rate
TR	Transfer payments to private households
VE	National income
w	Wage rate
WA	Employable population
WO	Total population
X	Real exports
Y	Real final demand
Y^*	Potential output
YV	Disposable income of private households

If the exchange rate against the mark, the anchor currency, is described by e, the price deflator of domestic demand by p and the price deflator of domestic demand in Germany by pf (all in natural logarithms), the domestic short-term interest rate by rs, the German short-term interest rate by rsf and the exchange rate risk premium by ri, then the following applies:

$$e = \gamma e_{-1} + (1-\gamma) f\left(rsf - rs, \frac{p}{pf} \right) \tag{1}$$

$$rs = \gamma(rsf + \Delta e + ri) + (1-\gamma) f(\ldots) \tag{2}$$

Under a strict fixed exchange rate system with $\gamma = 1$ and $\Delta e = 0$ equations (1) and (2) reduce to:

$$e = e_{-1} \tag{3}$$

$$rs = rsf + ri \tag{4}$$

$$ri = \beta ri_{-1} \quad \text{with} \quad \beta < 1 \tag{5}$$

Short-term interest rates in the ERM-countries have to be used to target the exchange rate against the mark. Therefore, they equal German short-term interest rates plus a risk premium which will be zero in the long run. In a system of fully flexible exchange rates with $\gamma = 0$, the exchange rates follow short-term interest rate differentials (uncovered interest parity) and in the long run purchasing power parity. In this case the short-term interest rate can be used as the monetary policy instrument to target monetary growth or the inflation rate. If a country abandons the obligation to intervene in the exchange rate mechanism it regains full influence on its short-term interest rate.

In reality the European Monetary System combines elements of a fixed exchange rate system with elements of a flexible exchange rate system. In the ERM, exchange rates can fluctuate in a margin around central rates which, in August 1993, was widened from ± 2¼% to ± 15%. Moreover, the central rates can be changed by realignments. Therefore, the policy parameter γ has to be set between zero (fully flexible exchange rates) and one (strictly fixed exchange rates). The exact value determines the respective changes in short-term interest rates and exchange rates.

Changes in short-term interest rates induced by monetary policy either in Germany or in other ERM-countries lead to changes in long-term interest rates and exchange rates. These in turn influence consumption, investment, exports and imports and therefore real GDP. Labour demand and the output gap depend on variations in real GDP. They again exert their influence on wages, production costs and domestic prices. Wages and employment determine labour income which in turn influences private consumption. Domestic prices as well as price expectations feed back into long-term interest rates, exchange rates and foreign trade.[9]

One of the main advantages of a structural macro-econometric model can be seen in the possibility to identify single channels in the monetary transmission process theoretically. The effects of a change in monetary policy on different sectors of the economy, e.g. on the consumption of private households or on the investment of private firms or on foreign trade, can be quantified empirically. The relative importance of the different transmission channels can be assessed.[10] Moreover the interaction of monetary and fiscal policy as well as changes in the policy-mix can be analysed.

The single country models have been specified in very similar but not fully identical ways. Structural differences between countries, especially in the financial sector of the economies and in the transmission process of monetary policy, are partly reflected in differences between the estimated coefficients but to some extent also in differences in the dynamic specification of the behavioural equations. Therefore, differences in the size and in the dynamics of the effects of a change in monetary policy between countries appear.

2. Convergence of short-term interest rates to a single monetary policy in EMU

With the prospective start of EMU in 1999, the European Central Bank will conduct a single monetary policy in the Euro area and set a common short-term interest rate level in the participating member states. Up to now, different interest rate levels in the selected European countries have prevailed. Short-term interest rates in Germany, the Netherlands and Belgium have already converged towards the lowest interest rate level in Europe for the last two years. The French short-term interest rate differential vis-à-vis these three countries was about two percentage points in 1995, but has narrowed strongly in 1996. The United Kingdom and Italy have higher and more persistent risk premia, with Italy's short-term interest rate being on the very top of these European interest rates. Nevertheless, in Italy, market expectations of participation in EMU from its start have allowed monetary the authorities to cut short-term interest rates substantially during the second half of 1996. With regard to interest rate differentials, it is an open question how interest rates will behave during the transition period to EMU and afterwards. A further interesting feature of the EMU process

9 Herrmann and Jahnke (1994).

10 Jahnke and Reimers (1995).

consists in the macro-economic effects of the convergence path towards a common EMU interest rate level. Some of these questions should be answered by the following simulation experiments which have been undertaken with the macro-econometric multi-country model of the Deutsche Bundesbank.

In order to calculate the effects of a single monetary policy in EMU, a baseline scenario has been designed. In this scenario, EMU will not start in 1999. Instead, the current ERM operates without tensions during the whole simulation period (from 1997 until 2003) and exchange rates between the mark and the other participating European currencies are stable. The German short-term interest rate follows an autoregressive path, while interest rates in the other countries are linked to the German short-term interest rate via a risk premium. According to this path, short-term interest rates will increase between 1997 and 2000, decrease thereafter and converge in the long run towards constant values which are 6½% p.a. for Germany, Belgium and the Netherlands, 7% p. a. for France, 9% p. a. for the United Kingdom and 12% p. a. for Italy. For the Netherlands and Belgium the risk premium is held at zero. Risk premia in France, the United Kingdom and Italy are positive, reflecting the interest rate differentials in 1996. To isolate the effects of EMU in the countries considered it is necessary to assume the participation of the United Kingdom in the ERM in this simulation. Therefore, the United Kingdom's entry into ERM hypothetically has been set on 1st January 1997, which seems to be early enough to fulfil the Maastricht exchange rate criterion.

The results of the baseline serve as a basis of comparison for two different policy simulation experiments, both assuming the start of EMU at the beginning of 1999. In the first experiment, short-term interest rates in EMU will converge towards German interest rate baseline levels, i.e. in the EMU area the relatively low German interest rate level will prevail. Since an over-night adjustment of interest rates on New Year's Eve 1998 seems not to be very realistic, a gradual path of convergence has been assumed during the years 1997 and 1998. This means that the positive risk premia of the pound sterling, the French franc and the Italian lira have been reduced exogenously every quarter by the same amount. After 1999 risk premia are zero for all participating currencies.

In the second simulation experiment, interest rates converge towards an average European interest rate level, calculated on the average of short-term interest rates in the baseline weighted by the respective GDP shares of the year 1995. Because of high interest rates in the United Kingdom and, in particular, in Italy the level of the average European interest rate is about 130 basis points above the German level.[11] As in the first case, interest rates converge gradually towards the common EMU level and risk premia are zero after 1999, too, with all these assumptions fully credible in the markets.[12] Regarding short-term interest rates in the United States, Japan and Canada, the assumption is made that they will not be influenced by the transition to EMU. Therefore, in the simulation experiments their interest rates are kept at baseline levels. In addition, fiscal policies in all countries are unchanged in comparison to the baseline solution.[13]

Charts 1 and 2 report the short-term interest rate paths in the respective countries for the different scenarios. The adjustment of short-term interest rates to German interest rate levels implies for France, the United Kingdom and Italy an expansionary monetary policy shock. The short-term interest rate in France falls by 50 basis points up to the start of EMU and in the United Kingdom it falls by 230 basis points. In Italy, the cut in short-term interest rate amounts to more than 500 basis points. In Germany, the Netherlands and Belgium, interest rates remain at baseline levels. The scenario with the adjustment of short-term interest rates to an average European interest rate level

---

11 This tightening of monetary conditions can be seen as a strategy of the national central banks and the ECB later on, to build credibility for the new monetary authority in Europe.

12 Similar interest rate scenarios have been described in Scheide and Solveen (1997).

13 Since the aim of this study is to isolate the effects of short term interest rate convergence, the consequences of the fulfillment of the convergence criteria, especially the fiscal criteria, have been excluded. In this respect, see e.g. Hughes Hallett and McAdam (1996) and Gros (1996).

Chart 1
Short-term interest rates

______ Base line
------- Adjustment to German interest rate level in EMU
.... Adjustment to average European interest rate level in EMU

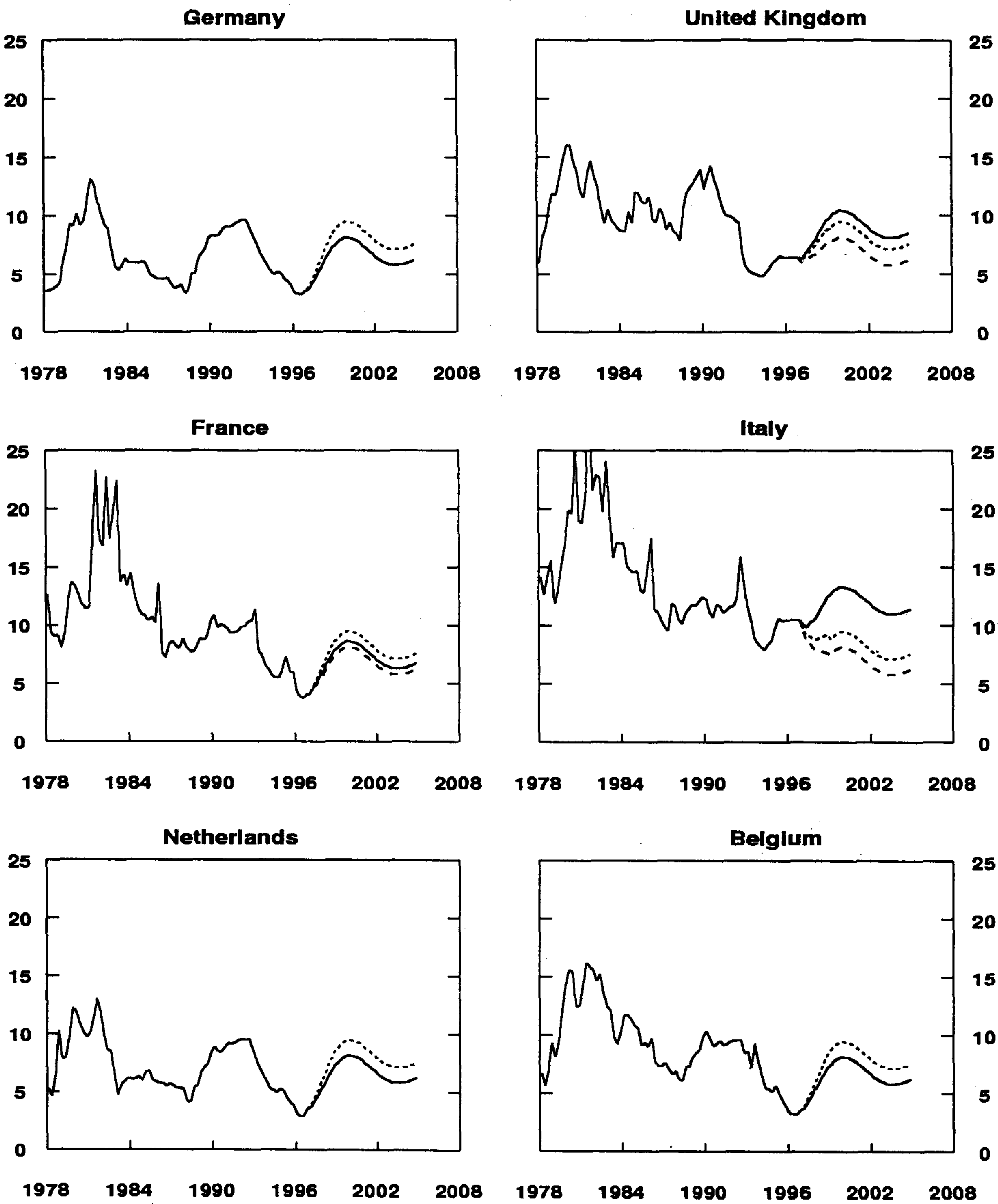
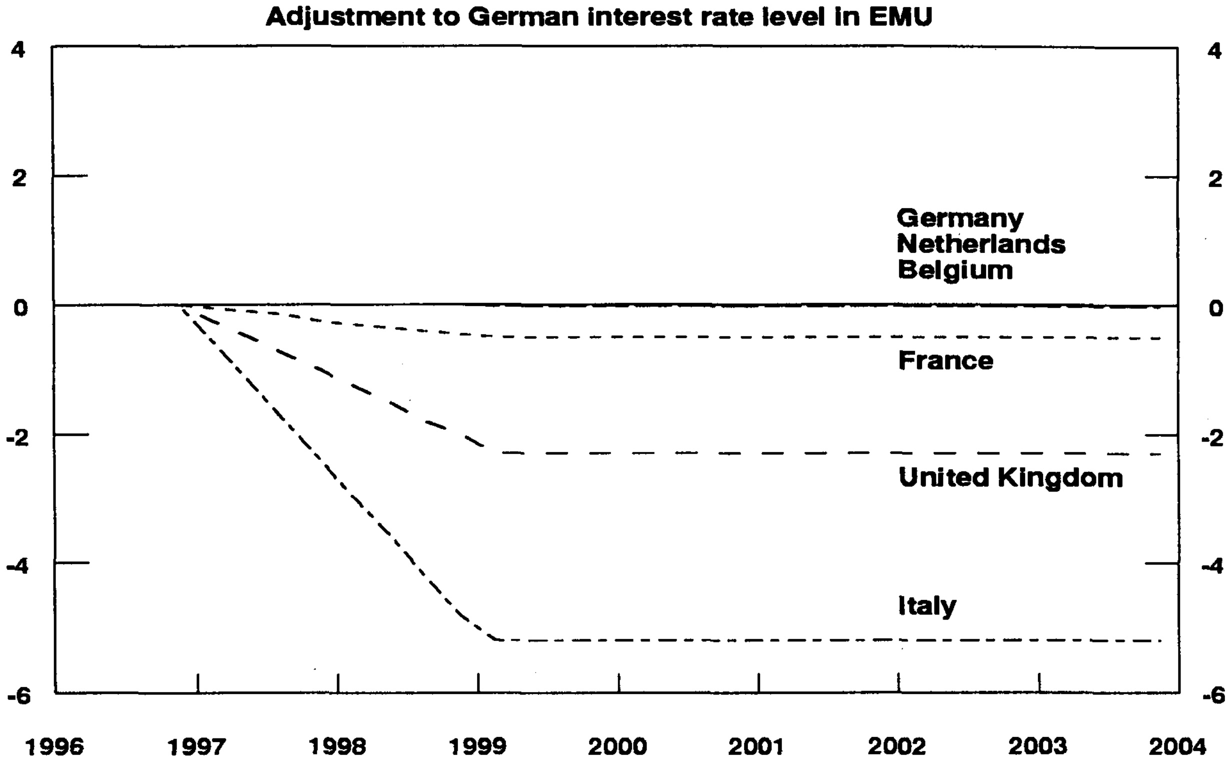
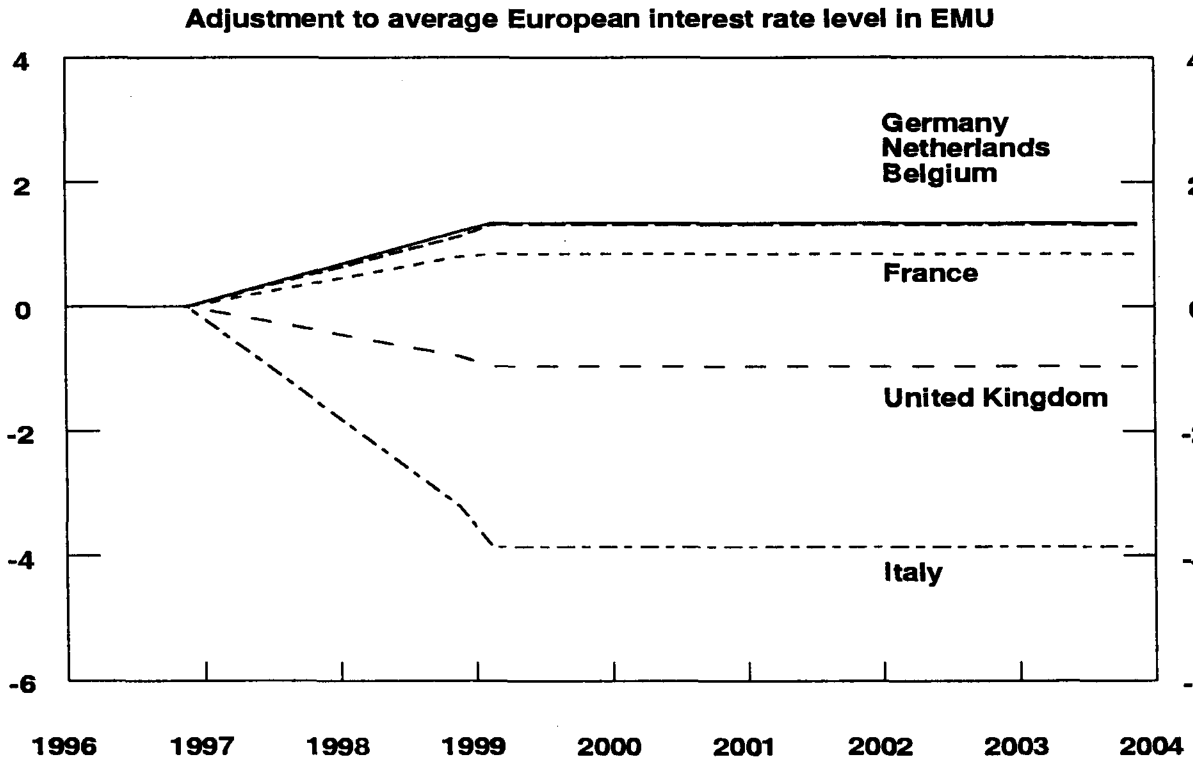


Chart 2

Effects of a single European monetary policy on short-term interest rates





involves for four countries a policy of monetary restraint. In Germany, the Netherlands and Belgium the rise in short-term interest rates is about 130 basis points altogether in 1999. The French interest rate increases by more than 80 basis points. At the same time, the United Kingdom and Italy still experience a loosening of monetary conditions with a reduction of short-term interest rates by roughly 100 basis points and 390 basis points respectively. It should be stressed that both simulation experiments imply a permanent change of interest rates. Fundamentally, this means a regime shift and may be subject to the Lucas critique. However, assuming that agents will not change their behaviour immediately, the application of the model to this kind of questions should still be admissible. But in the longer run, financial structures and transmission processes in the participating countries will adjust, at least to some extent.

3. Effects of a single monetary policy on output and prices

The results of the aforementioned changes in interest rates during the transition period to EMU and afterwards are described in Tables 3 and 4. Chart 3 depicts the response of real GDP to the two monetary policy scenarios during the simulation period. If interest rates in EMU adjust to the German short-term interest rate level all participating member countries will experience a rise in output due to decreasing real interest rates. As the graph in the upper part of Chart 3 shows, the expansionary effect of interest rate adjustment is, not surprisingly, strongest in Italy. After an interest rate cut by more than 500 basis points totally in 1999, output in Italy is 1.5% above the baseline level within five years of the start of interest rate convergence. In the United Kingdom, the peak effect is about 1.2% after five years, while in France it is above 0.2% within the same time period. The weaker effect can be explained by smaller interest rate cuts in these two countries compared to Italy.

On account of the international linkage considered in the multi-country model, output in Germany, the Netherlands and Belgium also increases in comparison to the baseline, though their short-term interest rates remain unchanged. In Germany real GDP is 0.2% above its baseline solution after five years. Remarkably, the effect on real GDP in the Netherlands and Belgium (0.5% and 0.3% respectively) is stronger than on the French real GDP. This is a reflection of the different degree of openness of the countries considered. In addition, different foreign trade elasticities also play a role. Looking on EMU aggregates (see Chart 5), the process towards a single monetary policy will have an expansionary impact of more than 0.6% on areawide activity if interest rates adjust to the low German level.

The lower part of Chart 3 plots the output reaction of short-term interest rate adjustments to an average European interest rate level. Here, the results are mixed. Recalling that this scenario means a tightening of monetary conditions in Germany, France, the Netherlands and Belgium, the reaction of real GDP in the respective countries is as expected. The gradual increase in Germany's short-term interest rates, amounting to 135 basis points altogether in 1999, leads to a reduction of German output by 0.35% in the same year; five years after the start of monetary tightening a negative deviation of Germany's real GDP by 0.1% still persists. In Belgium and especially in the Netherlands activity is much more reduced, though the interest rate shock in these countries is as high as in Germany. Again, the larger effect can be accounted for by the high degree of openness of these countries. In France too, activity is depressed by the higher interest rate.

Regarding the countries with a loosening of monetary conditions, the initial impact on growth is adverse in Italy because of the reduction in activity abroad. But in 1999, when the total interest rate cut of 390 basis points is operative, real GDP in Italy starts to rise as expected. A year later output is 0.4% above the baseline level, then falling again because of negative trade impacts. A remarkable case is the reaction of real GDP in the United Kingdom. Usually, one would expect a surge in activity if interest rates are cut by a total of almost 100 basis points. But in the United Kingdom output remains more or less at baseline levels, because the international interaction plays a role, too. Certainly, domestic activity in the United Kingdom is influenced positively through an interest rate reduction, but this reaction is not strong enough to overcompensate for the adverse effects on exports originating in the slowdown of activity abroad. For the EMU area as a whole, this scenario leads to a drop in real GDP by 0.15% after five years in comparison to the baseline (see Chart 5 and Table 4).

Table 3

Effects of a single European monetary policy in the multi-country model of the Deutsche Bundesbank: adjustment of short-term interest rates to German interest rate level

Deviation from baseline in percentages or in percentage points

	1997	1998	1999	2000	2001
1. Short-term interest rate					
Germany	0.00	0.00	0.00	0.00	0.00
United Kingdom	-0.64	-1.64	-2.30	-2.30	-2.30
France	-0.14	-0.38	-0.50	-0.50	-0.50
Italy	-1.50	-3.90	-5.20	-5.20	-5.20
Netherlands	0.00	0.00	-0.01	-0.01	-0.01
Belgium	0.00	0.00	-0.02	-0.02	-0.02
EMU	-0.38	-0.99	-1.34	-1.34	-1.34
2. Long-term interest rate					
Germany	0.00	-0.00	-0.00	-0.00	0.00
United Kingdom	-0.16	-0.47	-0.65	-0.50	-0.29
France	-0.01	-0.05	-0.06	-0.02	0.02
Italy	-0.44	-1.46	-2.06	-1.73	-1.18
Netherlands	0.00	-0.00	-0.00	-0.01	-0.00
Belgium	0.00	-0.00	-0.00	-0.01	-0.00
EMU	-0.10	-0.33	-0.46	-0.38	-0.24
3. Money stock					
Germany	0.00	0.03	0.10	0.21	0.33
United Kingdom	0.27	1.17	2.56	3.83	4.69
France	0.03	0.14	0.27	0.36	0.43
Italy	0.52	2.27	4.54	5.64	5.01
Netherlands	0.05	0.28	0.74	1.18	1.41
Belgium	0.03	0.16	0.45	0.80	1.09
EMU	0.14	0.63	1.39	2.04	2.36
4. Real GDP					
Germany	0.01	0.07	0.14	0.20	0.23
United Kingdom	0.11	0.44	0.85	1.09	1.16
France	0.01	0.07	0.16	0.21	0.23
Italy	0.08	0.46	1.06	1.48	1.53
Netherlands	0.02	0.14	0.30	0.41	0.46
Belgium	0.04	0.17	0.30	0.33	0.32
EMU	0.04	0.21	0.44	0.59	0.63
5. Private consumption deflator					
Germany	0.00	0.00	0.01	0.04	0.08
United Kingdom	0.01	0.07	0.29	0.70	1.18
France	0.00	0.00	0.02	0.06	0.13
Italy	0.01	0.13	0.50	1.12	1.80
Netherlands	0.00	0.00	0.02	0.07	0.15
Belgium	0.00	0.00	0.03	0.12	0.25
EMU	0.00	0.04	0.16	0.39	0.65

Table 4

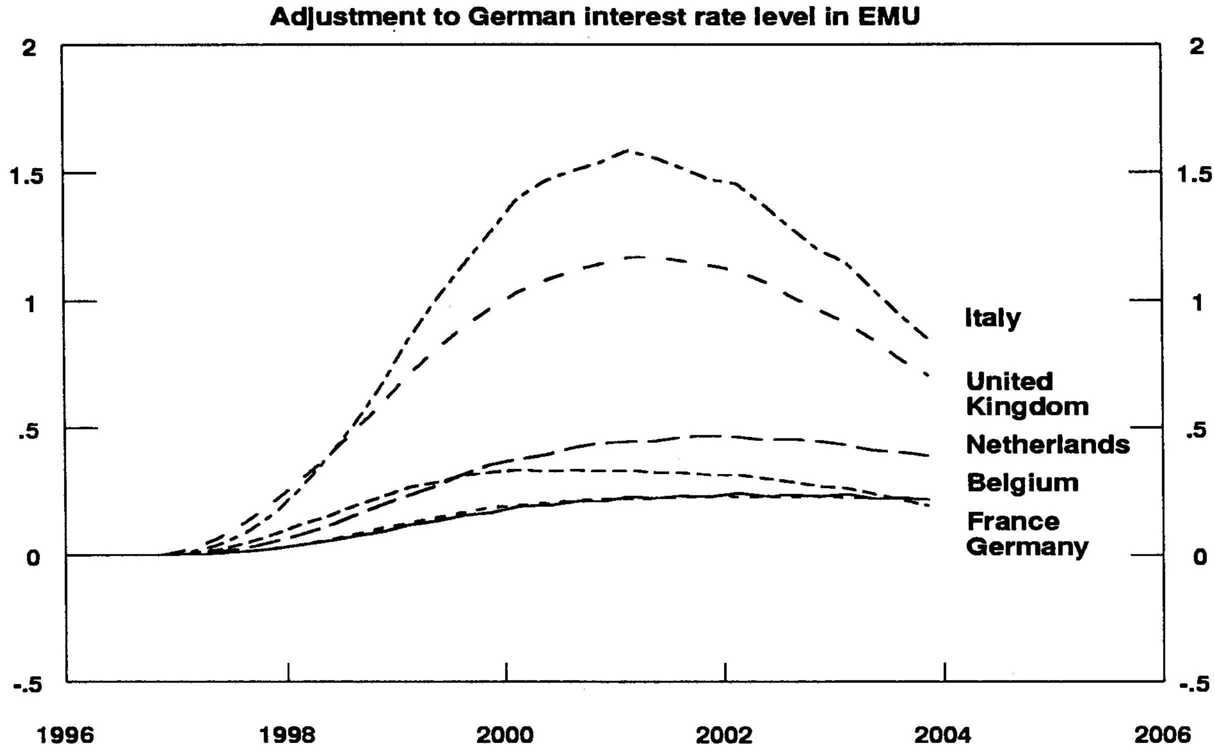
Effects of a single European monetary policy in the multi-country model of the Deutsche Bundesbank: adjustment of short-term interest rates to average European interest rate level

Deviation from baseline in percentages or in percentage points

	1997	1998	1999	2000	2001
1. Short-term interest rate					
Germany	0.38	0.97	1.34	1.34	1.34
United Kingdom	-0.25	-0.65	-0.96	-0.96	-0.96
France	0.25	0.65	0.84	0.84	0.84
Italy	-1.00	-2.60	-3.86	-3.86	-3.86
Netherlands	0.38	0.97	1.33	1.33	1.33
Belgium	0.35	0.91	1.32	1.32	1.32
EMU	0.02	0.05	-0.00	0.00	-0.00
2. Long-term interest rate					
Germany	0.08	0.25	0.35	0.26	0.13
United Kingdom	-0.06	-0.19	-0.31	-0.30	-0.26
France	0.03	0.08	0.08	0.00	-0.08
Italy	-0.30	-0.98	-1.55	-1.43	-1.10
Netherlands	0.13	0.44	0.64	0.56	0.41
Belgium	0.06	0.23	0.37	0.30	0.12
EMU	-0.01	-0.04	-0.10	-0.13	-0.16
3. Money stock					
Germany	-0.24	-0.82	-1.34	-1.39	-1.15
United Kingdom	-0.05	-0.05	0.03	0.07	-0.02
France	-0.17	-0.51	-0.77	-0.88	-1.01
Italy	0.19	0.89	1.93	2.35	1.73
Netherlands	-0.75	-2.54	-4.04	-4.08	-3.43
Belgium	-0.42	-1.62	-3.08	-3.84	-3.68
EMU	-0.16	-0.48	-0.68	-0.69	-0.70
4. Real GDP					
Germany	-0.12	-0.31	-0.35	-0.22	-0.13
United Kingdom	-0.02	-0.01	0.01	-0.01	-0.05
France	-0.08	-0.24	-0.32	-0.32	-0.36
Italy	-0.05	-0.02	0.19	0.39	0.38
Netherlands	-0.19	-0.53	-0.76	-0.83	-0.90
Belgium	-0.35	-0.71	-0.78	-0.69	-0.68
EMU	-0.10	-0.23	-0.24	-0.16	-0.15
5. Private consumption deflator					
Germany	-0.02	-0.14	-0.34	-0.57	-0.77
United Kingdom	-0.04	-0.17	-0.34	-0.45	-0.53
France	-0.02	-0.10	-0.23	-0.38	-0.56
Italy	-0.03	-0.14	-0.24	-0.21	-0.12
Netherlands	-0.01	-0.06	-0.18	-0.35	-0.56
Belgium	-0.01	-0.08	-0.29	-0.63	-1.02
EMU	-0.02	-0.13	-0.29	-0.45	-0.59

Chart 3

Effects of a single European monetary policy on real gross domestic product



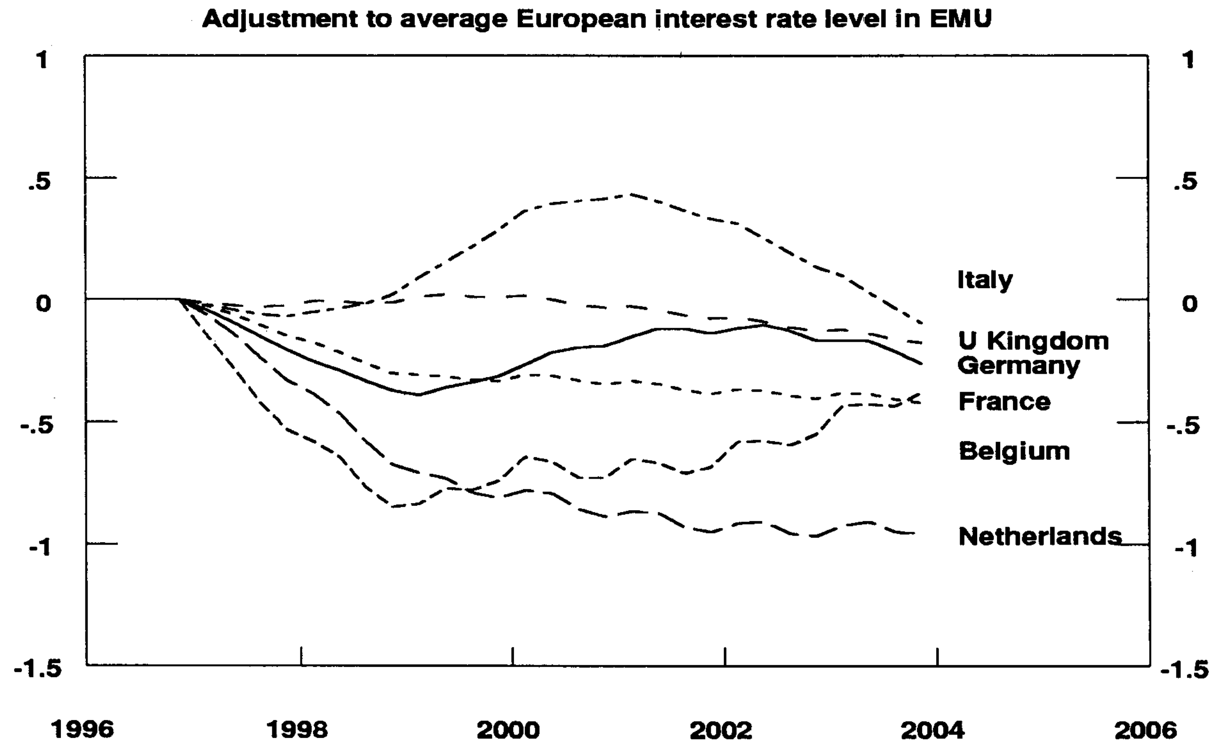
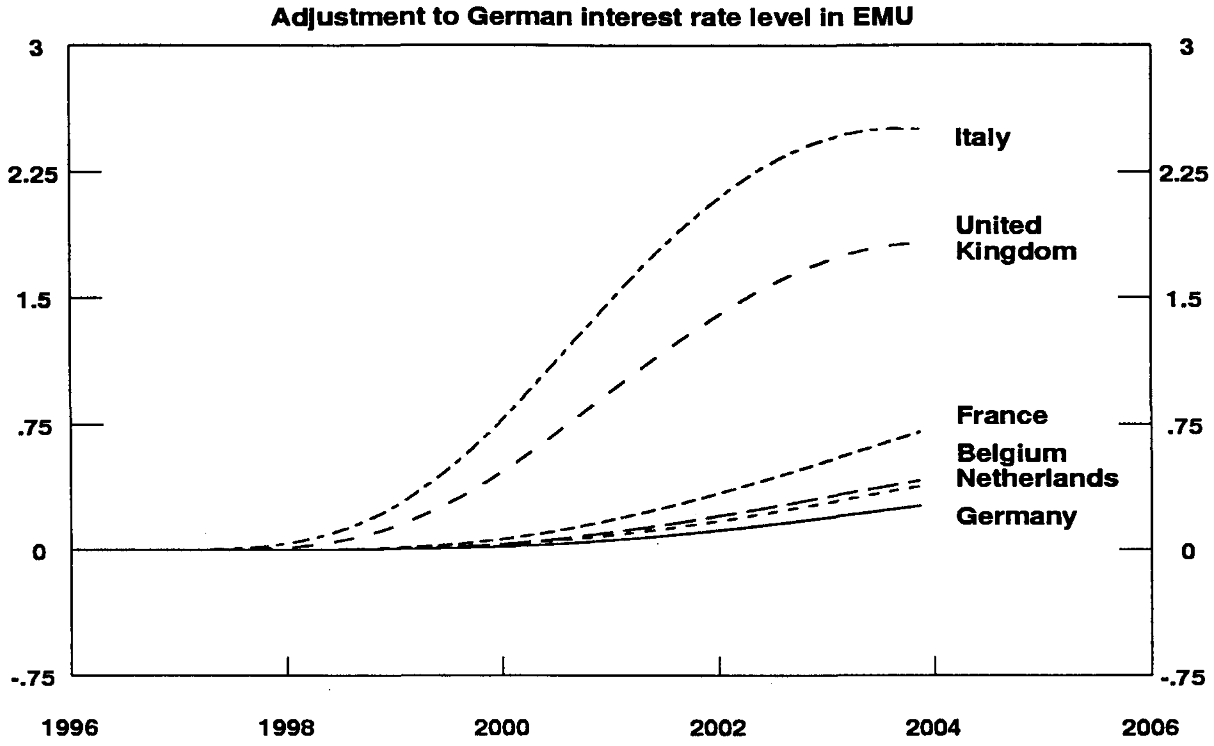


Chart 4
Effects of a single European monetary policy on private consumption deflators



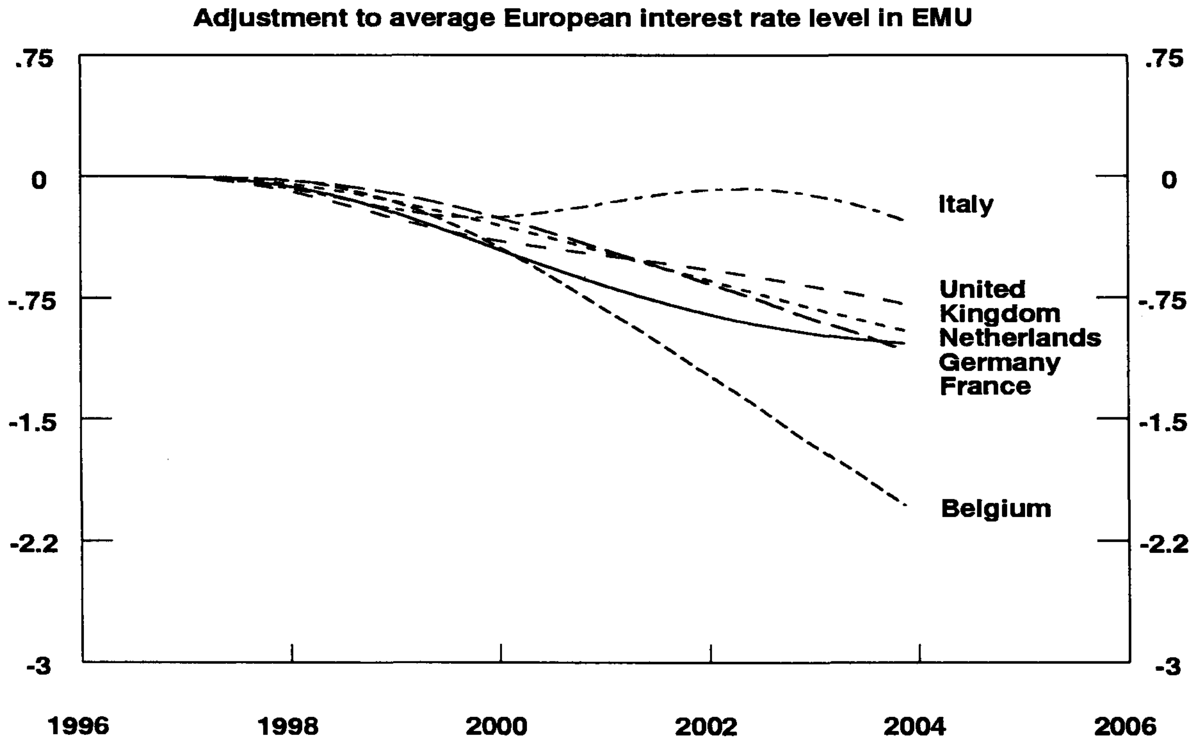
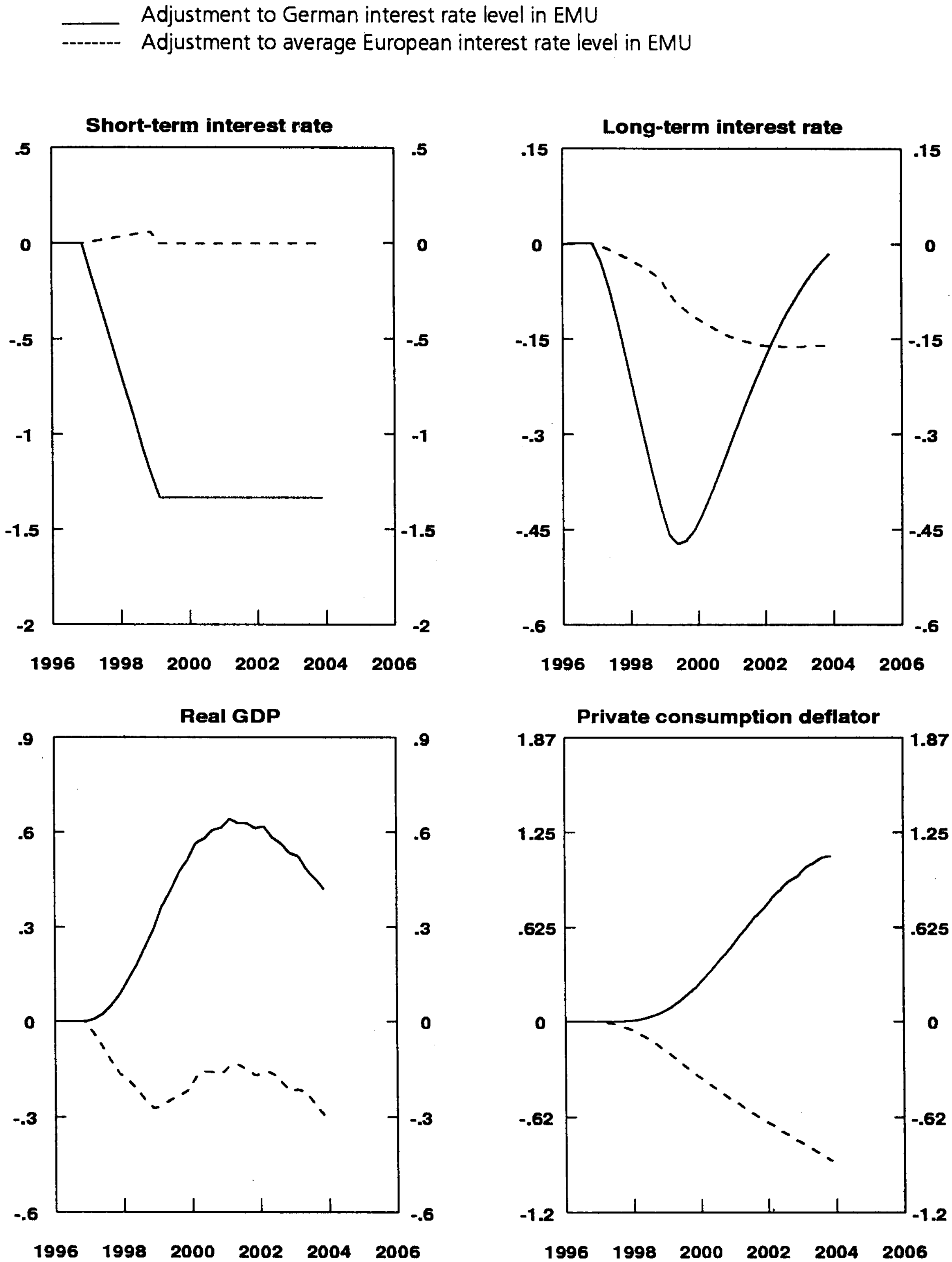


Chart 5

Effects of a single European monetary policy on EMU aggregates



The difference between the two simulation experiments is also evident in the simulated price responses. In the upper part of Chart 4 the effects of an interest rate adjustment to German interest rate levels on the private consumption deflators are depicted. As expected, prices rise in every EMU country, since the first scenario is expansionary. Corresponding to the size of the output effect the increase in consumer prices is the strongest in Italy. After five years the private consumption deflator in Italy is 1.8% above the baseline level, while still rising until 2003. The private consumption deflator in the United Kingdom shows a similar behaviour but on a lower scale. In Belgium, France, the Netherlands and, in particular in Germany, price effects are quite small initially. Five years after the start of interest rate convergence, Germany's private consumption deflator is less than 0.1% above its value in the baseline solution. In the later stages of the simulation period the increase in the price level is higher but still moderate. Aggregated over the whole EMU area the deflator of private consumption exceeds its baseline value by about 0.7%.

The price responses of interest rate adjustments to an average European interest rate level are reported in the lower part of Chart 4. Higher interest rates and a slowdown in activity in some countries lead to a fall in price levels in the whole EMU area. Surprisingly, even in Italy prices fall, though growth accelerates in this country. The reason for this unexpected development is decreasing import prices caused by falling export prices in the other EMU countries. Three years after the start of interest rate convergence the private consumption deflator in Italy is about 0.25% below baseline. Then the decreasing effect weakens and the price level deviation from the baseline fluctuates around zero. Since import price elasticities of domestic prices are very high in Belgium, the price drop is very strong in that country. Five years after the beginning of interest rate convergence the consumer price level in Belgium is about 1% below baseline. In the other countries, Germany, France, the Netherlands and the United Kingdom, price reactions are more moderate. In these countries the negative deviation of private consumption deflators range between 0.5% and 0.8% after five years. The aggregated EMU deflator of consumption is about 0.6% below its baseline level.

Conclusions

Even though many important issues on the way to EMU are still left open and will only be decided in the course of 1998, it seems necessary to assess the prospective effects of a transition towards a single European monetary policy. At the present time in January 1997, this is only possible under some very heroic assumptions regarding the participation of countries, the starting date and the conditions of entry into EMU. From different monetary policy scenarios which have been simulated with the macro-econometric multi-country model of the Deutsche Bundesbank only tentative conclusions can, therefore, be drawn.

If short-term interest rates in EMU converge towards the relatively low German rates, the effect on output in all participating countries will be expansionary, at least for some years. In this case inflation in EMU will accelerate and the price level will increase. The whole area gains in real output, but looses in price stability. This could endanger the credibility of the ECB and may lead to rising inflation expectations. If the European Central Bank adjusts its short-term interest rate to an average European level this will have an expansionary influence activity in some countries, mainly in Italy. But in other countries, including Belgium and the Netherlands, this policy will have contractionary output effects. This exerts a further dampening effect on prices even in those countries where the inflation rate is already very low. Output in the whole EMU area as well as the level of the private consumption deflator are below baseline levels in this case. In the transition period to a single monetary policy the whole area looses some real GDP, but gains in long-run price stability. The simulated reactions in output and prices depend, to some extent, on the expectation formation process used in the model and may be smaller and faster in the case of model-consistent expectations.

A single European monetary policy influences economic activity in the participating countries of EMU differently, at least in the transition period and in the first years of EMU, when structural differences between countries still exist. Real growth rates as well as inflation rates will, therefore, differ from country to country. According to the latest convergence report of the European Monetary Institute[14] the difference between the lowest inflation rate in Finland (0.9%) and the reference value (2.6%) in the period from October 1995 to September 1996 amounts to 1.7 percentage points. If such an inflation differential lasts for four years the price levels between different countries have already diverged by 7%. A single areawide monetary policy alone which targets aggregate goals does not seem to be suited to accelerate or even to maintain the convergence process between countries. On the contrary, the common policy may have divergent effects on the inflation rates in the participating countries. Instead, convergence has to be obtained by national economic policies, possibly in the field of fiscal policy and structural reforms to enhance flexibility and mobility in the labour, capital, goods and financial markets.

References

Bayoumi, T. (1996): "Who Needs Bands? Exchange Rate Policy Before EMU" in K. Alders, K. Koedijk, C. Kool and C. Winder (eds.) *Monetary Policy in a Converging Europe*, Dordrecht, Boston, London.

DeGrauwe, P., (1996): "How to Fix Conversion Rates at the Start of EMU". *CEPR Discussion Paper* No. 1530, November.

Deutsche Bundesbank (1996): "Makro-ökonometrisches Mehr-Länder-Modell", Frankfurt am Main, November.

European Monetary Institute (1996): "Progress Towards Convergence", Frankfurt am Main, November.

European Monetary Institute (1997): "The Single Monetary Policy in Stage Three – Specification of the Operational Framework", Frankfurt am Main, January.

Gros, D. (1996): "Towards economic and monetary union: Problems and prospects". *CEPS Paper* No. 65, Brussels.

Herrmann, H. and W. Jahnke (1994): "The interest rate policy transmission process in Germany". Bank for International Settlements, *National Differences in Interest Rate Transmission*, Basle.

Hughes Hallett, A. and P. McAdam (1996): "Fiscal deficit reductions in line with the Maastricht criteria for monetary union: An empirical analysis". *CEPR Discussion Paper* No. 1351, London.

Jahnke, W. and H.-E. Reimers (1995): "The transmission of monetary policy in the econometric model of the Deutsche Bundesbank for Germany". Bank for International Settlements, *Financial Structure and the Monetary Policy Transmission Mechanism*, Basle.

Jahnke, W. (1996): "Long-term interest rates and exchange rates in the Bundesbank macroeconometric model of the German economy". Bank for International Settlements, *The Determination of Long-term Interest Rates and Exchange Rates and the Role of Expectations*, Basle.

Scheide, J. and R. Solveen (1997): "Auf dem Weg zum Euro – Szenarien für Zinsen, Wechselkurse und Konjunktur", Kiel.

14 European Monetary Institute (1996).

Comments on : "Effects of a single European monetary policy: simulations with the multi-country model of the Deutsche Bundesbank" by Wilfried Jahnke and Bettina Landau

by C. E. V. Borio

The paper by Jahnke and Landau simulates the effects of the move to a common monetary policy associated with the establishment of EMU. The exercise is based on the Deutsche Bundesbank multi-country model and covers the period 1997-2003. It assumes a smooth transition: short-term interest rates converge in 1999 against the background of unchanged exchange rates among participating countries (Germany, France, the United Kingdom, Italy, the Netherlands and Belgium) and flexible rates vis-à-vis the rest of the world (specifically, the United States, Japan and Canada). Two scenarios are put forward. In the first, short-term interest rates converge to the level of those ruling in Germany. In the second, they converge to the average level prevailing in participating countries. The main results vary according to the scenario. In scenario 1 there is an overall expansion; given the differing starting points, this occurs at a different pace across countries. In scenario 2, there is an overall contraction and the behaviour of inflation and output diverges across countries. In other words, in this second case nominal convergence is associated with divergence in nominal and real variables.

This is a very neat and clear paper. It addresses an important question and provides some interesting answers. My comments will address essentially three questions. First, are the scenarios "reasonable" and properly handled by the model? By "reasonable" I do not mean "realistic": the paper is not attempting to provide a forecast; rather, the key criterion is internal consistency. Second, is the transmission mechanism implied by the model rich enough? Finally, what broader policy issues does the paper raise?

The scenarios and their handling

Much of what I will have to say on this relates to what is assumed to be exogenous in the simulations. The main point is that the assumptions about exogeneity may not be fully justified and can unnecessarily limit the validity of the results.

The paper posits an exogenous reduction in the "risk premia" on short-term interest rates. In fact, whether this can actually take place depends on other policies whose evolution is not consistently modelled. This is the case, most critically, for fiscal policy, both with regard to the Maastricht criteria for EMU eligibility and to the constraints imposed by the "Stability Pact" that would apply after 1999. Put differently, the "rules of the game" imply a specific relationship between the path of fiscal and monetary variables that would otherwise not exist. Some attempt to model this seems appropriate.

The paper does not really say much about the conditions under which each of the two scenarios would actually materialise and hence about what they would mean. In the model, convergence to the *average* level of interest rates (scenario 2) is an unambiguous tightening. In reality, this would presumably in part reflect less credibility on the part of the ECB compared with scenario 1 and thus higher long-term inflation expectations.[1] This would imply less of a tightening than assumed in the model. Simulations based on some form of model-consistent expectations would help here. They would seem more appropriate given the nature of the exercise.

---

[1] Of course, this would be different if higher short-term rates resulted from a pre-emptive strike by the ECB in an attempt to establish credibility.

The final point follows from the previous one. As posited in the present simulation, presumably in the scenario in which short-term rates converge to the average level there is a concomitant appreciation of European currencies vis-à-vis the other world currencies. This would strengthen the overall contraction. However, if the rise reflected an attempt to cope with a loss in credibility, this would not be the case. The observation reiterates the point that a careful use of model-consistent expectations – or at least ad hoc ways of capturing their effect – would have bigger implications than probably assumed by the authors.

The transmission mechanism

The paper does not explain much of the link between short and long-term interest rates. Given the major role that long-term rates play in the model, a richer discussion would be useful. This should cover not just changes but also the *levels* of the spreads, so as allow a better assessment of their plausibility. In fact, the model, as it stands, cannot account for one significant part of the story: for highly indebted governments default risk premia could rise as access to the printing press is centralised, tending to offset the reduction in the exchange rate (inflation) premium associated with a single monetary policy.

Similarly, it is not clear how much room the model allows for differences in countries' transmission mechanisms. A worthy example is the relative importance of short-term versus long-term rates. Previous studies, for instance, have clearly shown that in the United Kingdom and, to a lesser extent, in Italy, short-term rates play a more significant role than in the other continental European countries included in the model, where long-term rates are more influential.[2] This could well have an impact on the simulation results. One may wonder, for instance, if the surprisingly small increase in output in the United Kingdom to the 100 basis point cut in interest rates may not in part reflect the model restriction that monetary policy impulses are transmitted only via long-term rates.

Policy issues

It would have been useful had the authors expanded on the implications of their analysis for certain key policy questions touched upon in the conclusions. At least three sets of issues appear to be particularly relevant.

(i) What do the divergences they find for inflation and output imply for the credibility of the convergence process? And for the sustainability of EMU arrangements?

(ii) How strong will be the forces working towards a weakening of differences in the transmission mechanisms once EMU is put in place? How fast will those forces operate? Are the implications of such differences of first order significance?

(iii) What could the model tell us about the net effects of a weak and a strong Euro?

Clearly, these questions go well beyond the strict confines of the model. Yet the function of the simulation is precisely that of providing a benchmark for organising our thoughts and place us in a better position to find the corresponding answers.

---

2 See BIS (1995), *Financial structure and the monetary policy transmission mechanism*, CB 394, Basle, March.

Disinflation and credibility effects: the Swedish case[1]

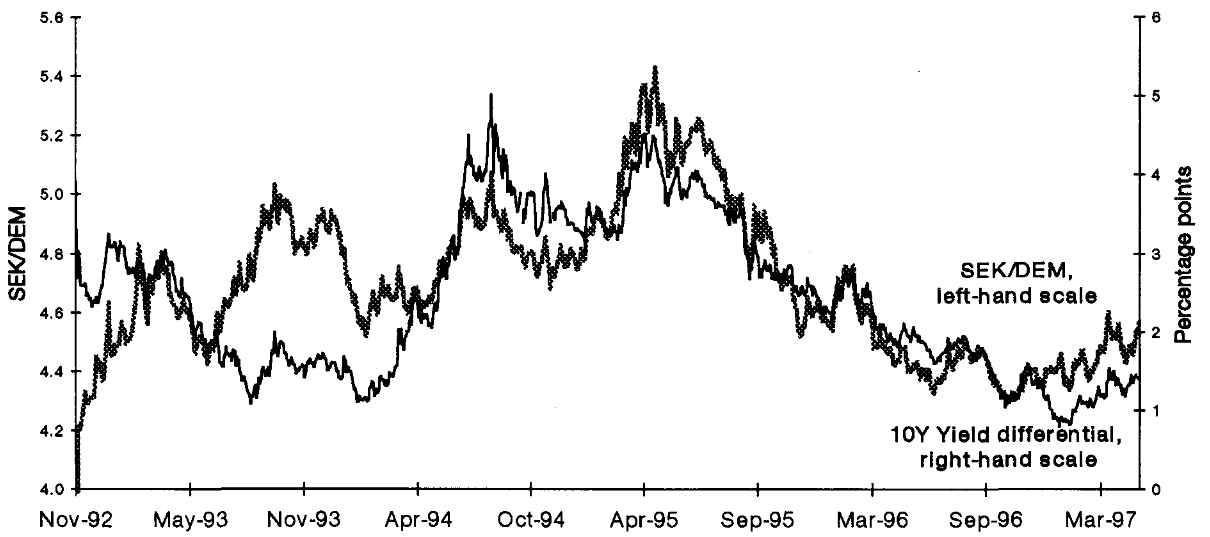
Hans Dillén, Tohmas Karlsson and Jonny Nilsson

Introduction

The interaction between financial markets in Sweden, on the one hand, and the real economy and the inflation process, on the other, has probably changed substantially in recent years due to the change from a fixed to a floating exchange rate regime with an explicit inflation target. There are several aspects of special importance. First, the adoption of an inflation target that differs markedly from the historical inflation rates in the seventies and eighties generated credibility effects that differed both in size and nature from effects associated with devaluation expectations in the fixed exchange rate regime. Moreover, uncertainties regarding the evolution of public finances have evidently reinforced these credibility effects.

The volatile developments in financial markets (see Figure 1) in recent years underline the importance of analysing the effects of imperfect and fluctuating credibility for economic policy, as well as the effects related to the monetary regime shift. It is not very useful to analyse the transmission mechanism without a framework in which the volatile developments of the exchange rate and long-term interest rate can be understood.

Figure 1
The (SEK/DEM) exchange rate and the long term yield differential



Apart from these credibility considerations, there are other changes that probably have altered the transmission mechanism per se. For instance, the role of the exchange rate channel is likely to be of greater importance in the present floating exchange rate regime. Moreover, the market's interpretation of monetary policy actions, such as changes in the repo rate (the key official interest rate), is different in the new regime with an explicit inflation target, as it modifies the transmission from the repo rate to market rates for longer maturities. These subjects, important as they are, will not be discussed in this paper.

Instead, this study tries to illustrate some of the consequences of the above-mentioned changes for the Swedish economy by means of simulations in a macro model. Especially, we will focus on the effects elicited by credibility shocks and the role of the disinflation process in this respect. We consider the main contribution of this study to be the work on modelling and quantifying credibility effects. We do not provide a detailed account of the macroeconomic development in Sweden since the announcement of the inflation target in January 1993. Instead, we regard this paper as an input to our ongoing calibration work on the model. At the present stage, we merely give a very stylised account of the *main* features of the Swedish disinflation process, in particular, the modelling of credibility effects and their repercussions on the real economy.

The paper is organised as follows. In Section 1, changes in the Swedish economy related to the monetary regime shift are discussed as well as different notions of imperfect credibility. In Section 2 we briefly discuss the main features of the macro model we employ for the simulation analysis.[2] Next, in Section 3 we present two alternative scenarios for a disinflation process, with different kind of credibility effects. In the final section, we summarise and draw conclusions.

1. The monetary regime shift and credibility effects in Sweden

1.1 The monetary regime shift[3]

Traditionally, a fixed exchange rate has served as an intermediate target for the ultimate objective of price stability in Sweden. After the break-down of the Bretton Woods system in 1972 Sweden participated in the so called currency snake, which can be seen as a precursor to the EMS. In August 1977 the krona was taken out of the snake system and unilaterally pegged to a weighted index, which in 1991 was replaced by the ECU. During the seventies and eighties, actual and expected inflation was, however, too high relative to other countries and the fixed exchange rate regime was sometimes not credible. This development resulted in several devaluations and, finally, an abandonment of the fixed exchange rate regime in November 1992. Against this background, in January 1993 the Governing Board of the Riksbank adopted an explicit inflation target as a means to guiding inflationary expectations and in this way improve the credibility of monetary policy. The inflation target was set to 2% with a tolerance interval of ± 1%, and it became operative in 1995.[4]

As will be discussed shortly, the credibility of the low inflation policy has improved substantially, even though this favourable trend periodically has been interrupted by events related to

---

2 We gratefully acknowledge the very generous support and help from the Bank of Canada provided to us. Without their assistance our macro model project would not have been possible.

3 For a more detailed discussion of the monetary regime shift in January 1993 and the consideration behind the introduction of the new inflation target in Sweden, see Andersson and Berg (1995) and Svensson (1995).

4 In this context, inflation is measured as the twelve-month increase in the consumer price index and the inflation target should be evaluated against the average inflation rate during a year. In 1993 and 1994, i.e. before the inflation target became operative, monetary policy was directed towards preventing the underlying rate of inflation to be rising. At the time the underlying rate of inflation was around 2%.

political uncertainty as well as to uncertainty regarding the development of the fiscal position in Sweden. The next section defines in more detail and discusses what is meant by imperfect credibility.

1.2 Credibility effects

Before making statements about how the credibility of an economic policy directed at price stability changes over time, it is necessary to define credibility and relate it to some quantitative measure. In general, credibility problems of monetary policy arise when expectations of the target variable (e.g. the exchange rate or the inflation rate) deviate from the declared target. However, inflation expectations are difficult to observe and interpret and similar problems arise when it comes to the definition and quantification of imperfect credibility. Moreover, even in an economy characterised by perfect credibility with respect to the inflation target, actual inflation as well as inflation expectations will deviate from the target due to unavoidable short-run inflation effects. It is, however, extremely difficult to separate these effects from those due to imperfect credibility. With these considerations in mind, the gap between inflation expectations and the inflation target will be viewed as an indicator rather than a perfect measure of credibility problems.

Furthermore, in analysing the economic effects that imperfect credibility may have we find it useful to distinguish between two types of credibility: operational and political.[5] Operational credibility refers to mechanisms that can be analysed endogenously within our model whereas political credibility refers to effects that have to be treated as exogenous shocks. It is worth emphasising that this decomposition of credibility into an endogenous and an exogenous part reflects features of the model used in this study and should not be seen as a universal modelling device.

Before discussing in more detail and separately how to model operational and political credibility we will present an overview of these concepts and point out some important features. Operational credibility mainly refers to the extent to which agents expect that future inflation will be in line with the inflation target *within the* current regime. In other words, operational credibility does not incorporate effects arising from expectations that the current target for monetary policy might change in the future. Operational credibility thus largely depends on the conduct of monetary policy and will improve if actual inflation evolves in accordance with the inflation target. Furthermore, the degree of operational credibility is mainly reflected in surveys of inflation expectations and it heavily influences price and wage setting behaviour but also to some extent financial prices, since changes in inflation expectations affect expectations of future monetary policy actions and thereby short term interest rates.

Political credibility, as used in this paper, reflects investors' fears that the economy will switch to an inflationary regime, which in turn will increase long-term interest rates and the exchange rate. The quantitative effect of such fears can be formalised as a regime shift premium representing the probability of a regime shift times the change in the average rate of inflation associated with this regime shift.[6] The effects of imperfect political credibility probably depend on the degree of political support for the inflation target and the perceived sustainability of public finances; experience tells us that these factors are of greater importance during periods of international turmoil on the financial markets. Political credibility is, therefore, treated exogenously because its determinants are very hard to model endogenously when long-run inflation expectations are tied to the inflation target and the sustainability of public finances is a constraint. As we will see in Section 1.4, shocks to political credibility will be modelled as exogenous, simultaneous shocks in long-term interest rates and the exchange rate in a manner that can be justified theoretically.

---

5 Operational and political credibility are concepts introduced and discussed by Andersson and Berg (1995).

6 See Dillén (1994, 1996) and Dillén and Hopkins (1997) for analysis of regime shift premia and their impact on interest rates. Regime shift premia in the exchange rate is analysed by Dillén and Lindberg (1997) and discussed in Section 1.4.

Table 1
Features of operational and political credibility

Type of credibility	Operational	Political
Determinants	Actual inflation and monetary policy actions.	Political support for the inflation target. The perceived sustainability of public finances. Actual inflation.
Characteristics	Events that cause a deviation between inflation expectations (within the current regime) and the inflation target.	Occasional and exogenous shocks often related to international and political events.
Endogenous or exogenous in this study	Endogenous.	Exogenous.
Indicators	Surveys of inflation expectations. Financial prices to some degree.	Financial prices. Surveys of inflation expectations to some degree.
Effects	Price and wage setting behaviour and to some extent financial prices.	Financial prices, especially long term interest rates and the exchange rate.

In practice it is often difficult to distinguish between operational and political credibility and the two mechanisms regularly interact. If actual inflation substantially exceeds the inflation target it will mainly damage operational credibility but political credibility may weaken as well. On the other hand, weak political credibility will give rise to volatile financial markets and therefore makes the transmission process more unpredictable. This in turn affects the conduct of monetary policy. A notable difference is that political credibility to a lesser extent is reflected in surveys of inflation expectations. Political credibility represents expectations of potential shifts to high inflation regimes, which should be included in inflation expectations. However, it seems that such expectations are, to only a limited extent, incorporated in surveys of inflation expectations. One explanation is that surveys reflect agents' assessment of the most likely outcome for future inflation rather than inflation expectations in a mathematical sense.[7]

1.3 Operational credibility and endogenous inflation expectations

When the Swedish inflation target was announced in 1993, inflation expectations tended to be above the target, especially the long-term expectations. Since then inflation expectations have been revised downwards to levels close to the inflation target.[8] This development is most pronounced for long-term inflation expectations, whereas short-term inflation expectations are also strongly affected by the actual evolution of inflation. Overall, there seems to be a gradual adjustment of inflation expectations towards the inflation target, which probably to a large degree can be attributed to improved operational credibility.

---

7 However, as discussed by Dillén and Hopkins (1997), it appears that surveys of financial investors' inflation expectations to some extent reflect regime shift expectations.

8 During the period 1993-95 one-year inflation expectations of households normally fluctuated in the upper half of the tolerance interval (1-3%) whereas expectations at or below the target have been observed since 1996. One-year inflation expectations in industry were even higher (3-4%) before they at the end of 1995 rapidly started to decline towards the lower limit of the tolerance interval. Bond investors' long-run inflation expectations have exhibited a declining trend from almost 5% in early 1993 towards slightly below 3% at the end of 1996. For further details about these surveys, see various issues of the Riksbank's Inflation Report.

Inflation expectations were normally above the actual inflation outcome during this period, indicating imperfect operational credibility. Inflation expectations that systematically exceed the actual inflation outcome do not in this case necessarily imply irrational expectations, but rather asymmetric information. Even if price stability is the overriding objective of monetary policy, this is not known by private agents and the degree of commitment to the announced target has to be learned by examining the actual developments of inflation and how monetary policy responds to different shocks. This gradual adjustment of expectations can be modelled within the model in a straightforward way. The formation of inflation expectations can schematically be described as:[9]

$$\pi^e = \alpha\pi_{backward} + (1-\alpha)\pi_{perfect\,foresight} \tag{1}$$

As seen from equation (1), inflation expectations combine a backward-looking component, given a weight α, with a forward-looking (or model-consistent) component, given a weight $(1-\alpha)$. Inflation expectations represent perfect foresight if $\alpha = 0$. The backward-looking component can be seen as adaptive expectations that reflect the historical pattern of inflation. If historical inflation is significantly above the target when this is announced, inflation expectations will be above the target. However, if actual inflation evolves in accordance with the target then this backward-looking component will converge towards the target. In this way inflation expectations will gradually adjust towards the target. The formation of expectations is endogenously determined within the model, implying that the analysis of operational credibility problems, as indicated by the gap between expectations and the target, are treated *endogenously.*

1.4 Political credibility effects as exogenous shocks to financial markets

Expectations and credibility problems are in practice complicated concepts and very difficult to characterise fully endogenously. The long-run credibility of the inflation target is mainly affected by mechanisms (often of a political nature) that are difficult to capture within the model. In particular, this type of imperfect credibility is suitably modelled as expectations of shifts to high inflationary regimes (see e.g. Dillén (1994, 1996) and Dillén and Hopkins (1997)). Such regime shift expectations are not easily incorporated into our model, which assumes only one regime with a single-steady state inflation rate. In a fixed exchange rate regime, similar credibility problems in the form of devaluation expectations will arise. As pointed out by Dillén (1994), a loss of exchange rate target credibility mainly increases short-term interest rates, whereas a loss of inflation target credibility is likely to increase long-term interest rates.[10] To formalise these ideas we consider the following expression for the logarithm of the nominal exchange rate:

$$s = s_0 + s_{rs} \tag{2}$$

where s_0 is the normal component of the exchange rate, reflecting normal factors affecting the exchange rate including normal monetary policy actions, and s_{rs} is the regime shift component, reflecting the effects of monetary regime shifts expectations. For example, Bertola and Svensson (1993) derive an expression for the exchange rate essentially of the form of equation (2) in which the regime shift component is proportional to the expected rate of devaluation.[11] In the case of a free floating exchange rate regime with an inflation target, Dillén and Lindberg (1997) derive a

---

9 A more detailed description of how expectations are formed within the model will be presented in Section 3.

10 In the regime switching model of Dillén (1994) a loss of credibility of an inflation target means a larger probability of switching to a high inflation regime.

11 See Bertola and Svensson (1993, equation (19)), in which the regime shift component is of the form $s_{rs} = \alpha g$, where g is the expected rate of devaluation. Notice, however, that g also appears in some exponential terms that take care of the so called "smooth pasting conditions".

closed-form expression for the regime shift component representing the effects of investors' fears that the economy will shift to a high inflation regime. Moreover, given that Germany is a country with very high credibility with respect to its low inflation policy, it can be shown that the regime shift component is approximately related to the long term yield differential, d^L, relative to Germany, as:

$$s_{rs} = \gamma d^L \tag{3}$$

One way to understand equation (3) is to consider the degree of monetary policy credibility as a demand factor. A reduction of the credibility of an exchange rate or inflation target will lower the demand for bonds denominated in the domestic currency, which presumably will lead to higher interest rates as well as a weaker currency. A fluctuating regime shift premium can, in this way, explain a positive correlation between the exchange rate and the interest rate differential. Conceptually, regime shift premia represent peso type credibility problems, since shifts to a more inflationary regime typically do not occur in the analysis. On a more general level, however, equation (3) can be seen as representing a fluctuating demand factor that can incorporate risk premia of various kinds. In what follows we therefore include the possibility that the regime shift premium also incorporates other kinds of premia that give rise to similar effects.[12]

In order to get a quantitative feeling for the credibility effects, let us reconsider Figure 1. During 1993 the long-term interest rate (differential) gradually declined from the very high levels that prevailed during the currency crises in 1992; to some extent, this can be seen as a gradual adjustment of expectations towards a low inflation regime. The depreciation of the krona during 1993 is somewhat hard to explain but, on occasions, political credibility shocks seem to be present. From 1994 to 1996 there is clear evidence of political credibility shocks of the kind discussed above. First, notice the positive correlation between the long-term interest rate differential and the exchange rate – a phenomenon that is difficult to explain without introducing a fluctuating regime shift premium (or another demand factor). It is also difficult to attribute the large movements in the exchange rate and the interest rate differential to other economic factors. Thus, it seems that a fluctuating regime shift premium might have been the dominant determinant of the volatile evolution of the exchange rate and the interest rate differential in recent years. Given this conjecture, the value of parameter γ in equation (3) can be assessed by "eye-econometrics": the fall of the long-term interest rate differential by slightly more than 3% since April 1995 has been accompanied by an appreciation of around 15%, which suggests a γ-value in the range 4 – 5. However, more careful analyses that take other important determinants into account sometimes suggest a lower value for γ. In the simulations we assume a γ-value of 4.

2. Presentation of the simulation model [13]

The model used in this paper consists of two sub-models: the *"steady-state model"*, which describes the long-run equilibrium features of the economy, and a *dynamic model* that explains short and medium-run adjustment paths of relevant macro variables when, after shocks, the economy is off its steady-state path. Permanent shocks will by definition change the steady state.

---

12 Due to peso type problems it is very difficult to distinguish between regime shift premia and other kinds of premia with similar characteristics. Any kind of a statistical verification of a regime shift model requires observations of regime shift. However, during the period with a fixed exchange rate policy several regime shifts in the form of devaluations were observed in Sweden.

13 Our current model is a slightly modified version of the Quarterly Projection Model (QPM) of the Bank of Canada. For a full documentation of QPM see Black, Laxton, Rose and Tetlow (1994) and Coletti, Hunt, Rose and Tetlow (1996).

The model assumes an exogenously growing one good economy characterised by a demographic structure with overlapping generations and non-Ricardian features. The model describes the behaviour of households, firms, the government, the central bank and the rest of the world. The decisions of these agents interact to determine the ultimate levels of key stocks: capital, government debt, and net foreign assets. These target stock levels in turn are key determinants of the associated flows, such as consumption, saving, investment, government spending and revenues, and the external balance. There is a formal stock-flow accounting framework that ensures full consistency among all variables both in the long run and along the dynamic adjustment path.

The model provides solutions for both the desired wealth of consumers in the long run and the consumption/savings paths that will sustain that level. Household supply of labour is treated as exogenous. Firms take the real cost of capital as given and choose the optimal stock of capital to go with it, as well as the path for investment spending that will take the economy to that equilibrium and maintain it. The government chooses a steady-state ratio of government debt to GDP. With these three steady-state decisions taken, aggregate net borrowing or lending for the economy as a whole is determined, resulting in the net foreign asset position of the country. Associated with this equilibrium net foreign asset position will be a unique external balance, i.e. imports, exports and foreign debt service. The relative price that will adjust to achieve this is the real exchange rate.

An important characteristic of the model is that the steady-state allocation of real variables is independent of the target rate of inflation. This implies, among other things, that the long-run level of all real stocks are left unchanged. Accordingly, there is no mechanism in the model that determines an optimal rate of inflation. That has to be determined by factors outside the model and then imposed as the inflation target rate of the monetary authorities.

The dynamics in the model originate from mechanisms explaining *gradual adjustments* of real as well as nominal variables such as prices and wages. The dynamics stem from three different sources. The first source is intrinsic to the economic structure and refers to all sources of gradual adjustment not related to expectations. These include labour market contracts, the fixed costs associated with investment, and so on. Such features give rise to a gradual response to disturbances, regardless of how large the disequilibrium might be. One can think of this as a general phenomenon of costly adjustment, which causes all agents in the economy to choose not to adjust immediately to disturbances.

A second source of dynamics is the *expectation mechanisms* in the model. Expectations are formed in a flexible way. In principle they are constructed as a combination of *backward- (adaptive) and forward-looking* elements. The weights attached to the two types of expectations differ between variables. A high degree of backward-looking expectations implies a high degree of rigidity in the adjustment of a variable. For instance, it takes time before changes in monetary policy feed into price and wage formation, which in turn implies a need for substantial adjustments in real variables. In contrast, a high degree of forward-looking expectations generates more flexible adjustments.

The third source of dynamics is the reaction of policies to disturbances. As the main purpose of the model is to analyse monetary policy, the endogenous reaction of the fiscal authority is not modelled in all its details. In order to respect the governments budget constraint and targets for public expenditure and debt as ratios to GDP, the income tax rate changes endogenously. Furthermore, the model includes a monetary policy reaction function, according to which a rise in anticipated inflation above target produces a rise in interest rates (slope of the yield curve) intended to move inflation back towards its target level over a horizon of six to seven quarters. The horizon is not arbitrary but an approximation of the sort of horizon over which monetary policy can have a meaningful influence on the trend rate of inflation. With monetary policy represented in this way, the response on the part of monetary authorities to economic disturbances that affect inflation is built-in; i.e. interest rates (slope of the yield curve) adjust automatically to put inflation on a path that will converge towards its target level.

More formally, the dynamic elements give rise to the following general structure of the determination of the GDP-deflator:

$$LP_t = \alpha LPBACK_t + (1-\alpha)LPFOR_t + \phi MAV(GAP_t) + \upsilon MAV(GAPPOS_t) \tag{4}$$

$$LPBACK_t = A(L)LP_{t-1} \tag{5}$$

$$LPFOR_t = \sum_{s=0}^{T} \rho^s E_t\left(LP^*_{t+s}\right), \qquad 0 < \rho < 1 \tag{6}$$

$$E_t(LP_{t+s}) = \delta B(L)LP_{t-1} + (1-\delta)LP_{t+s} \tag{7}$$

$$R^s - R^L = \theta\left[\tfrac{1}{2}\sum_{s=6}^{7}\left(\pi_{t+s} - \pi^{target}\right)\right] \tag{8}$$

where *A(L)* and *B(L)* are lag polynomials in *LP*. Equation (4) shows that the (log of) the GDP deflator has a backward and a forward-looking part. Behind this general formulation lies an intertemporal pricing problem for a firm facing adjustment costs in changing its price. The backward-looking part, equation (5), is derived from the fact that the *rate of change*, not just the level, is of importance. The forward-looking part, equation (6), consists of the discounted expected future *desired* price levels, with the horizon truncated *T* periods ahead. The desired price level is given by the static first order condition for the firm, taken from the steady-state model. Expectations of the future price level are modelled as a mixture of adaptive and model consistent elements according to equation (7).

Equation (8) describes the monetary policy rule. If actual inflation, six to seven quarters into the future, exceeds (is lower than) the target rate of inflation, the monetary authority raises (reduces) the short term interest rate. Eventually the inflation gap is eliminated through movements in the slope of the yield curve, short-term interest rate minus long-term rates, which in turn induces an output gap, *GAP*. As can be seen from equation (4), the output gap enters into the determination of prices in an asymmetric way. A positive output gap will have a larger inflationary impact on prices compared to the deflationary effects that a corresponding negative gap gives rise to. This general model structure of the dynamic behaviour of the GDP deflator is common to nominal wages as well as to the deflators of the GDP-components.

The determination of the short and the long-term interest rate and the nominal exchange rate is highly endogenous. The interdependence goes via the yield curve and the assumption of uncovered interest parity (UIP). However, expectations about the future nominal exchange rate do not exhibit perfect foresight. There are some expectation errors in the short run. In the long run the nominal exchange rate is determined by the level of the real exchange rate and the difference between the domestic and the foreign price level. Of these only the real exchange rate is solved in the steady-state model, since there is no price level targeting. The actual levels for prices will depend on the exact type of shock under study and the dynamic factors that drive the inflationary process until the target inflation rate is reached again.

The long-term interest rate is partly determined by the expectation theory of the term structure and partly by the inflation differential with rest of the world. Furthermore, there is a direct link between shifts in the short-term rate and movements at the longer end of the yield curve, assumed to represent the empirical observation of a high degree of volatility in long-term interest rates in Sweden; i.e. more than the pure expectation theory predicts. In the long run, interest rates and the nominal exchange rate obey relative PPP; i.e. they adjust to accommodate differences in actual and expected rates of inflation between Sweden and the rest of the world.

3. Simulation results

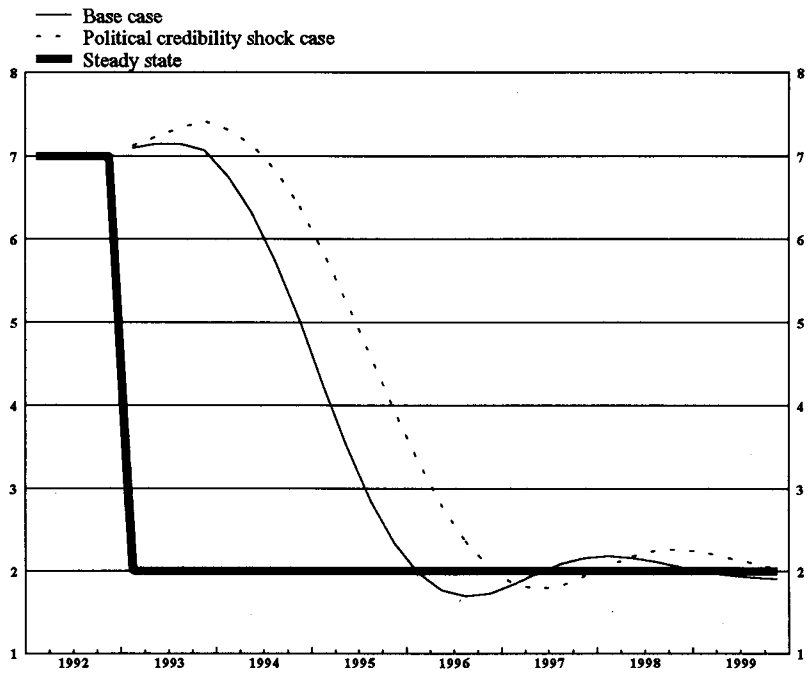
The purpose of this section is to illustrate some of the main effects of introducing an inflation target when the credibility of this target is imperfect. Since our main focus is not on tracking historical data very closely, we make the simplifying assumption that we start our simulations from a steady-state situation. A more realistic analysis would start from a disequilibrium situation that resembles the Swedish macroeconomic environment at the time of the introduction of the inflation target. Instead, we try to implement the concepts of operational, or endogenous, credibility and political, or exogenous, credibility in the context of a general equilibrium model.

Section 3.1 presents the general characteristics of two different disinflation scenarios and relates the model simulations to operational and political credibility. Next, in Section 3.2 we discuss in more detail the results when only operational credibility effects are taken into account. This scenario is taken as our base case. In Section 3.3 the outcome of a scenario containing operational as well as political credibility effects is compared with our base case.

3.1 Two disinflation scenarios

Our base case disinflation scenario entails a permanent reduction of the target rate of inflation from 7 to 2% per annum. The reduction of the domestic inflation rate by 5% illustrates the shift from a high inflation regime during the eighties, with an average inflation rate of about 7%, to the present low inflation regime where the official target rate of inflation is 2%. In addition, we assume that the foreign target or long-run inflation rate is reduced from 4 – 2% annually.

Figure 2: Annual inflation rate (%)



In Figure 2, the thick line shows the shift in steady-state or target inflation from 7 to 2% taking place in the first quarter of 1993. The thin continuous line represents the transition path for inflation during our base case scenario. Disregard the dotted line for the time being. Even in the base case scenario it takes time to change expectations of private agents about future inflation, since expectations are modelled as a weighted average of backward and forward-looking elements, as can be seen from equation (1) above. Consequently, the adjustment to the new and lower inflation target will be a gradual process.[14] Figure 2 shows that it takes approximately three years to reach the new target rate of inflation.

The time it takes to reach the new inflation target is partly dependent upon the relative importance of the backward-looking component of expectations (α in equation (1)). The larger this weight, the more importance agents attach to the past when forming expectations about future inflation. If we assume completely model consistent expectations ($\alpha = 0$), the history of inflation plays no role at all in forming inflation expectations. In this special case, agents fully believe the new inflation target in the sense that past inflation outcomes are not considered. In general, however, agents pay attention to previous inflation outcomes when forming expectations about future inflation ($0 < \alpha < 1$). Assuming that the central bank does its job, actual inflation will gradually approach the target rate and this in turn causes inflation expectations to be revised downwards towards the new target rate. This feature of expectations formation is *operational* or *endogenous credibility.*

In the second scenario, we attempt to model *political* or *exogenous credibility* effects. Informal empirical evidence suggests that a 1% increase in the long-term interest rate differential is associated with a 4% depreciation of the exchange rate as described by equation (3) above. We take this observation as a guide to the determination of the magnitude and correlation of exogenous shocks to the nominal exchange rate and the domestic long-term interest rate.

There is an instantaneous 4% depreciation of the Swedish krona which abates over four quarters and then reverses somewhat for another four quarters; this pattern is consistent with the depreciation of the krona in the period following the introduction of the inflation target. This disturbance is accompanied by a rise in the long-term interest rate differential towards the rest of the world.

Our second disinflation scenario contains two elements of credibility. First, we have operational or endogenous credibility through the gradual revision of expectations. Second, we have added an exogenous component, political or exogenous credibility, by shocking the nominal exchange rate and the domestic long-term interest rate.

3.2 Disinflation with purely endogenous credibility effects

In this section we give an account of the disinflation process when only endogenous credibility effects are present – our base case. The results are presented in Figures 3-18 below. The dotted line in these figures, representing the political credibility shock case, will be analysed in the following section.

Generally, monetary policy affects the rest of the economy via three channels: (i) a change in interest rates; (ii) real exchange rate movements through the uncovered interest rate parity condition; and (iii) altered expectations. Immediately after the announcement of the new inflation target, an inflation gap arises; i.e. a deviation between the actual and the target inflation rate, amounting to 5%, as can be seen from Figure 3.

To reduce the inflation gap, short-term interest rates have to increase in order to restrain aggregate demand and enforce the new inflation target on the expectations of private agents. The short-term interest rate is raised by roughly 500 basis points (Figure 5), whereas the long-term rate actually falls somewhat due to lower expected future inflation, as shown in Figure 6. The slope of the yield curve, i.e. the short minus the long rate, is shown in Figure 4.

---

[14] However, due to adjustment costs for real as well as nominal variables, there is a gradual response to shocks even in the absence of partly backward-looking expectations.

Figure 3: Inflation gap (shock - control, %p.)

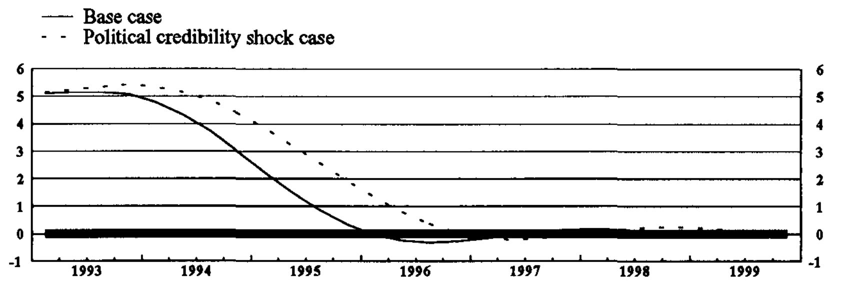


Figure 4: Short minus long interest rate (shock - control, b.p.)

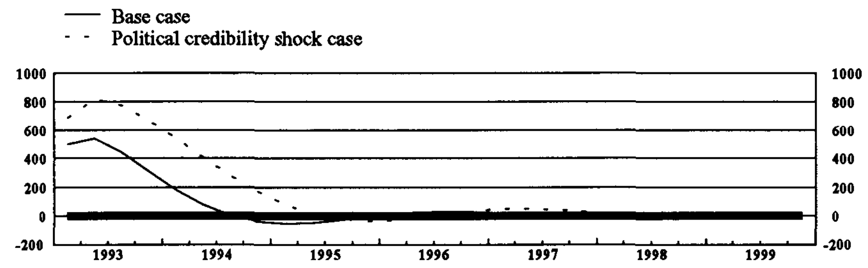


Figure 5: Short term nominal interest rate (shock - control, b.p.)

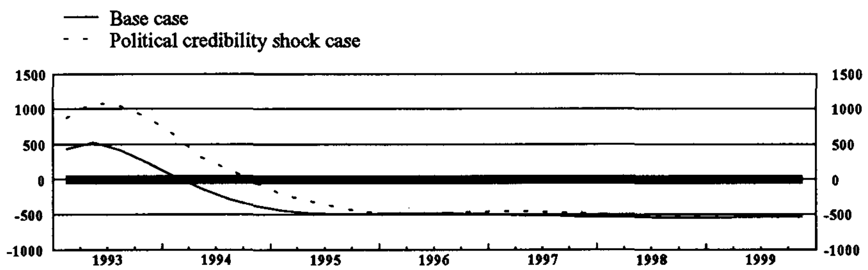


Figure 6: Long term nominal interest rate (shock - control, b.p.)

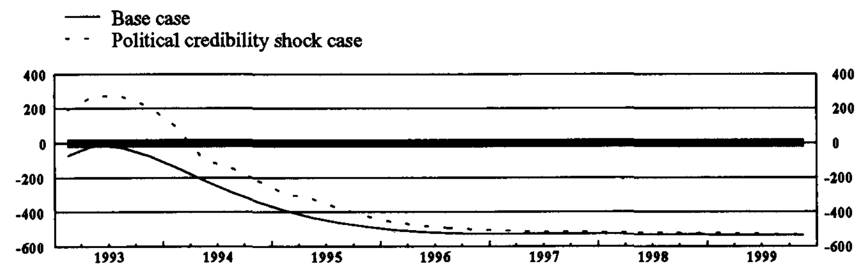


Figure 7: Output gap (shock - control, %)

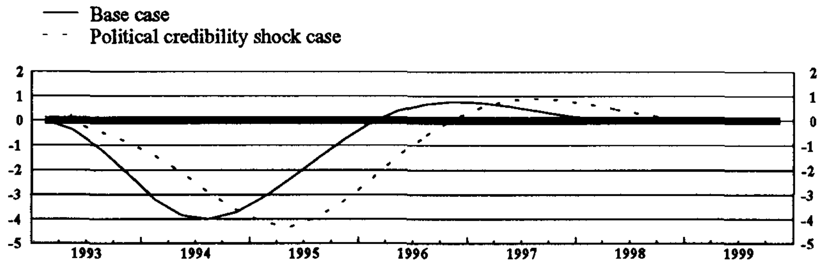


Figure 8: Cumulative output loss (shock - control, %)

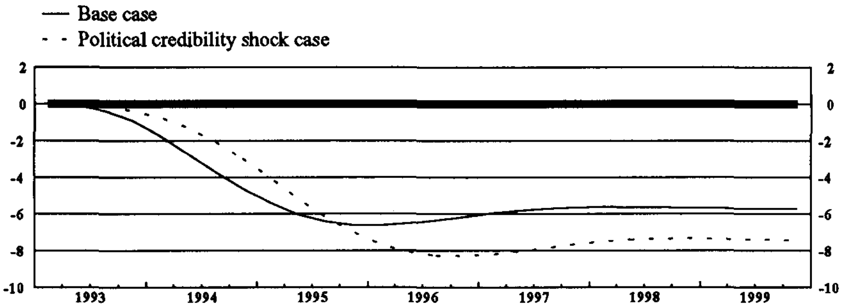


Figure 9: Consumption (shock - control, %)

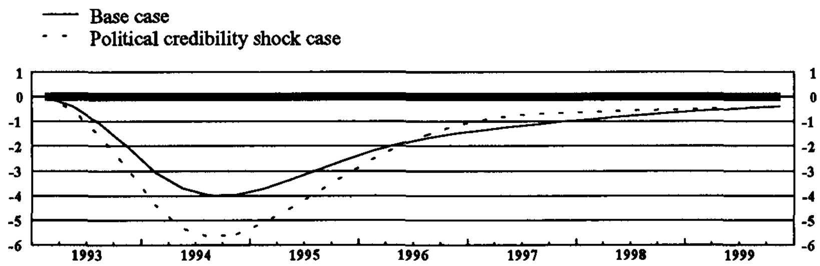


Figure 10: Investment (shock - control, %)

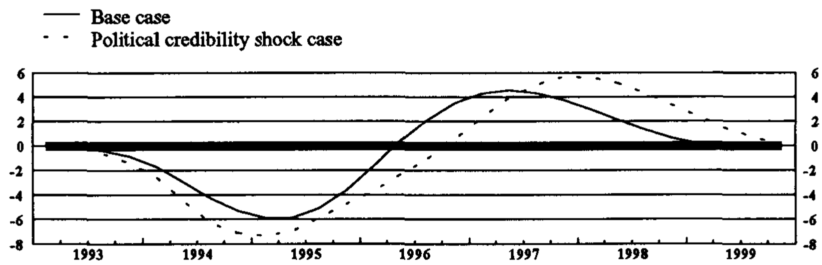


Figure 11: Net exports (shock - control, %)

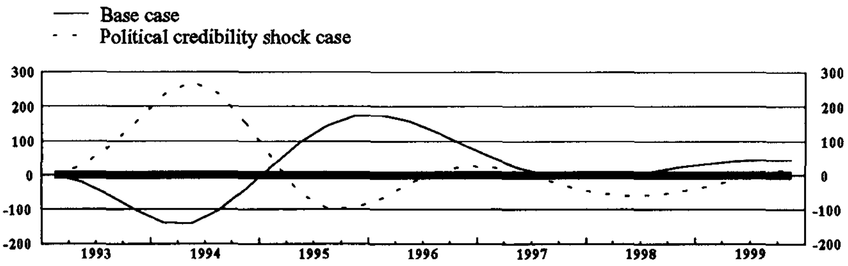


Figure 12: Real exchange rate (shock - control, %)

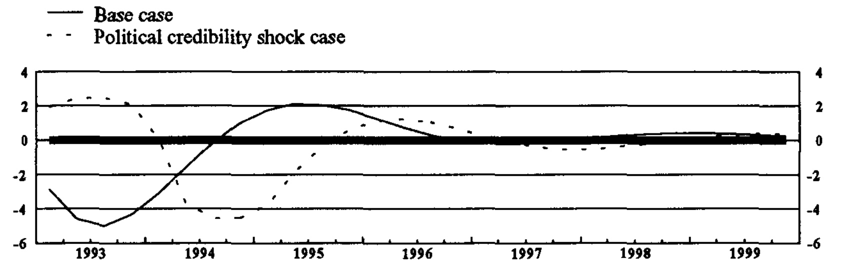


Figure 13: Exports (shock - control, %)

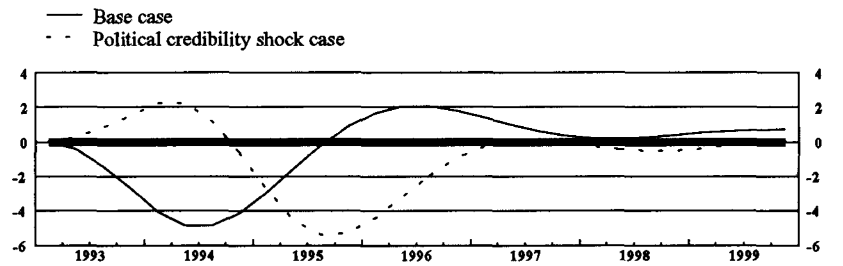


Figure 14: Imports (shock - control, %)

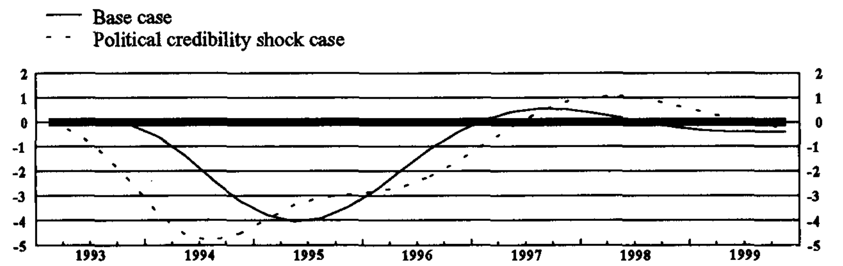


Figure 15: Nominal exchange rate (index)

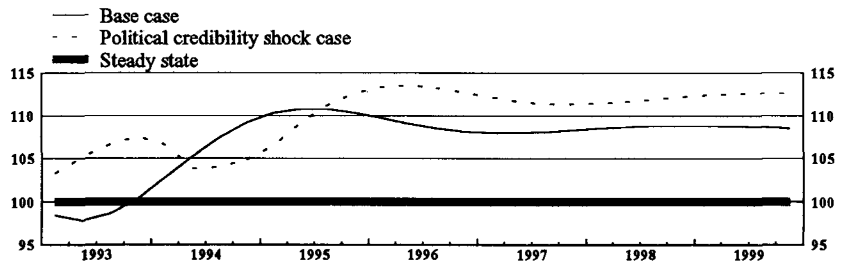
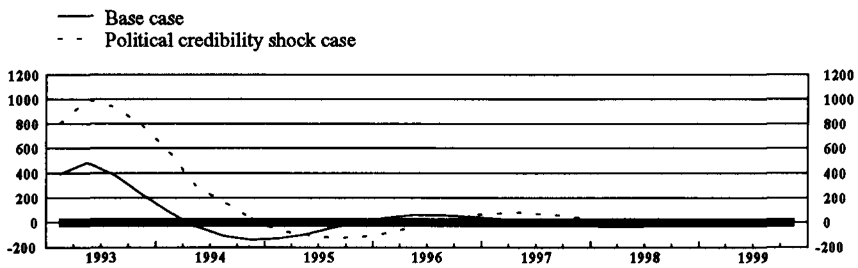


Figure 16: Real short term interest rate (shock - control, b.p.)



The tight policy pursued by the monetary authorities reduces aggregate demand and its components and creates a negative output gap (Figure 7), that reaches a minimum of -4% of GDP after roughly a year and a half. This negative output gap in turn has a dampening effect on prices and wages via the Phillips curve. Figure 8 shows the accumulated output loss associated with a monetary policy pursued in order to get inflation back on target.

The slope of the yield curve, expected income and changes in wealth determine private consumption in the short and medium run. All three components act to reduce consumption (Figure 9), by approximately 4% after six to seven quarters. Investment is mainly driven by the real cost of capital and an accelerator mechanism, both of which have a dampening effect, as is evident from Figure 10. The real cost of capital depends primarily on the real interest rate, which rises due to the monetary policy response as seen in Figure 16.

In addition to the interest rate channel, the transmission of monetary policy works through the exchange rate. The rise in real interest rates creates a slight appreciation of the real exchange rate through the uncovered interest parity condition (Figure 15). Given that prices are sticky in the short run, the real exchange rate also appreciates, as shown in Figure 12, which in turn causes net exports to fall, as seen from Figure 11.

The separate effects on exports and imports are depicted in Figures 13-14. The dampening of imports through the decline in consumption and investment is offset during the first year by a real appreciation of the krona such that imports are more or less unchanged during this time.

Eventually, all variables will settle down to their previous steady-state values, except for nominal variables such as the nominal exchange rate, which exhibits a drift in its level as witnessed from Figure 15. This ends our tour of the base case disinflation scenario. We will next discuss the effects of incorporating a political credibility shock.

3.3 Disinflation including political credibility shocks

The same basic mechanisms that were at work in the base case disinflation scenario are also operative in our alternative scenario incorporating political credibility shocks. However, the introduction of political credibility shocks elicits some additional effects.

First of all, the initial depreciation gives rise to an additional inflationary impulse, as can be seen by comparing the two inflation paths in Figure 3. Via the UIP condition the current nominal exchange rate is connected to its expected value the next quarter. The expected nominal exchange rate depends, in turn, on expected prices and the expected real exchange rate. The price level drift that accompanies the propagation of the initial shock will give rise to further depreciation tendencies and thereby prolong inflationary pressures. In order to defend its inflation target, the monetary authority is forced to raise the short-term interest rate considerably more than in the base case. Figure 5 shows that the difference in response is a hefty 500 basis points. In order to interpret this figure correctly one has to remember that we want to reduce the inflation rate by 5%, starting from a steady-state situation with full employment of all resources. Long-term interest rates also rise substantially, partly in response to higher expected future inflation but also as a direct consequence of the long-term interest rate shock (Figure 6).

From Figure 4 we note that the monetary policy response, as measured by the spread between the short and long-term interest rates, is severely tightened compared with the base case. This has obvious effects on output, consumption and investment. The negative output gap is prolonged, compared with the base case, see Figures 7-8. Another difference is the development of foreign trade. Due to the initial depreciation shock the resulting weak currency provides a stimulus to net exports. Compared with the base case, we tend to get a "dual economy", in which the traded goods sector is stimulated through a temporarily depreciated real exchange rate, whereas the non-traded or sheltered sector is depressed by the tight monetary policy stance (Figures 9-11).

Inflation settles down to the new target about one year later and, in the meantime, inflation is higher compared with the base case. Thus, the upward drift in the price level and the nominal exchange rate is more pronounced, indicated by an extra 4% depreciation of the nominal exchange rate in the long run, as shown in Figure 15.

A measure of the real cost imposed by the political credibility shock, on top of what is already inherent in the base case, is given by the difference in the cumulative output loss associated with the two scenarios. As is evident from Figure 8, we now have an accumulated output loss that amounts to about 8% of annual GDP, compared with 6% in the base case. Thus, the additional real cost amounts to roughly 2% of GDP. Although one should bear in mind the highly stylised character of our simulations, the results indicate that political credibility effects are non-trivial for reasonable magnitudes of the imposed exogenous shocks.

3.4 A closer look at expectations

Since expectations, especially those concerning inflation, play an important part in the disinflation story we would like to discuss the effects on inflation expectations per se, in addition to what is already incorporated in the behaviour of other macro variables. We compare actual inflation, one-year and two-year expected inflation in Figures 17-18 below. The one and two-year expected inflation rates are defined as follows:

$$E_{t+i}(\pi_{t+i+4}) = \left[E_{t+i}(LP_{t+i+4}) - E_{t+i-4}(LP_{t+i})\right] \quad i = 0,\ldots,n \tag{9}$$

$$E_{t+i}(\pi_{t+i+8}) = \left[E_{t+i}(LP_{t+i+8}) - E_{t+i-4}(LP_{t+i+4})\right] \quad i = 0,\ldots,n \tag{10}$$

where i=0 denotes the current period and n is the length of the simulation period.

Actual inflation vs. one-year inflation expectations

Figure 17A: Base case (%)

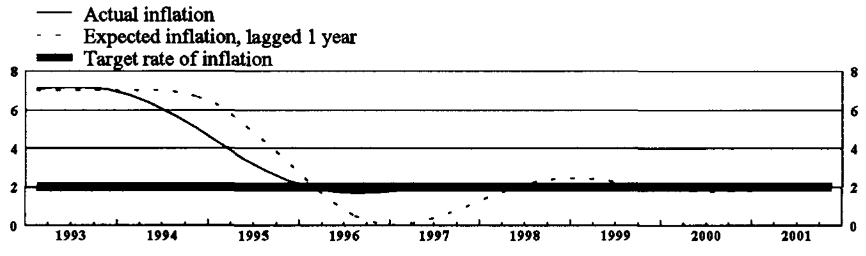
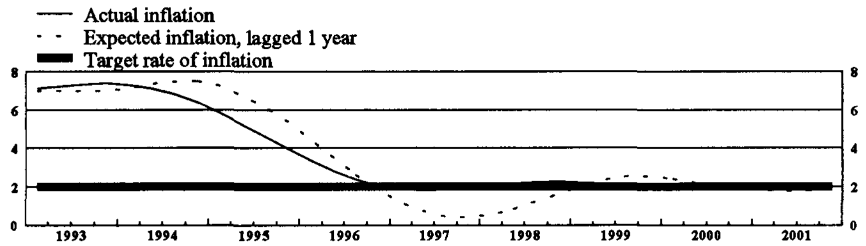


Figure 17B: Political credibility shock (%)



In the base case (Figure 17A), we compare the actual inflation rate during the disinflation process, denoted by a continuous line, with the inflation rate expected one year ahead, represented by a dotted line. In order to facilitate the comparison, the dotted line shows the expected one-year inflation rate lagged one year, which means that for each date on the horizontal axis the actual and the expected inflation rate refer to the same period. Hence the vertical difference between the dotted and the continuous line can be interpreted as a measure of the expectation error for a given period. For instance, in the period 1995Q1 in Figure 17A, the actual inflation rate is approximately 4%, whereas the expected inflation rate, close to 6%, is considerably higher giving an expectation error that overstates actual inflation by roughly 2%.

As would be expected from partly backward-looking expectations, expected inflation lies above actual inflation for the period immediately following the regime shift (1993Q1), as is evident from Figure 17A. However, from the beginning of 1996 until the middle of 1998, expectations underestimate the target inflation rate, sometimes quite considerably. Figure 17C shows the difference between expected and actual inflation; i.e. the expectation error for a one-year horizon in our base case. From Figure 17A we notice that actual inflation is inside the tolerance interval, $\pm 1\%$, of the target around the middle of 1995. This date is indicated by a vertical line in Figure 17C. Deviations of expected from actual inflation, after that period, can be interpreted as a measure or indication of a lack of credibility of the inflation target.

Whereas it seems quite reasonable that it takes some time before inflation expectations are revised downwards when the new and lower target is introduced, we are less happy about the subsequent rather large and persistent undershooting of inflation expectations. In Figure 17C it takes almost five years before expectations finally settle down inside the tolerance interval. Although there is only very scant empirical evidence regarding expectations, a cursory look at the available survey

Figure 17C: One-year inflation expectation error

Base case (%p.)

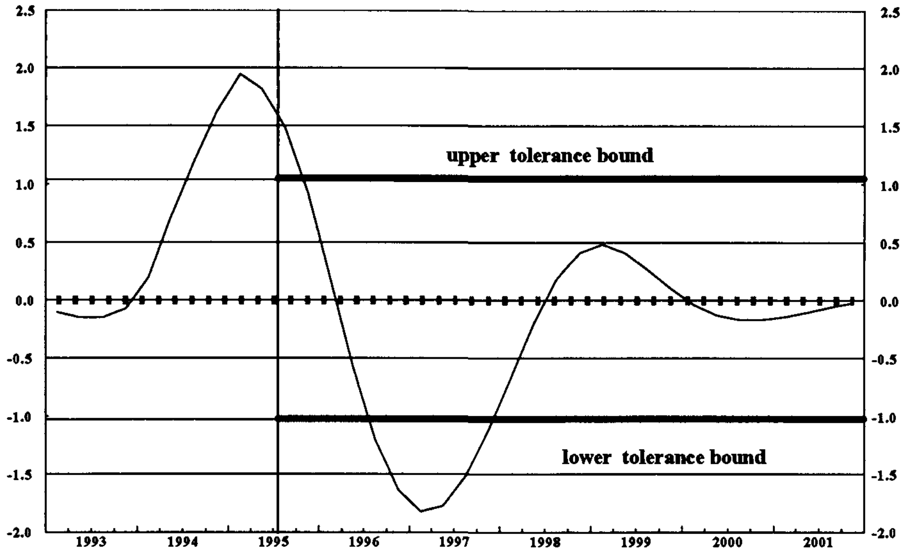
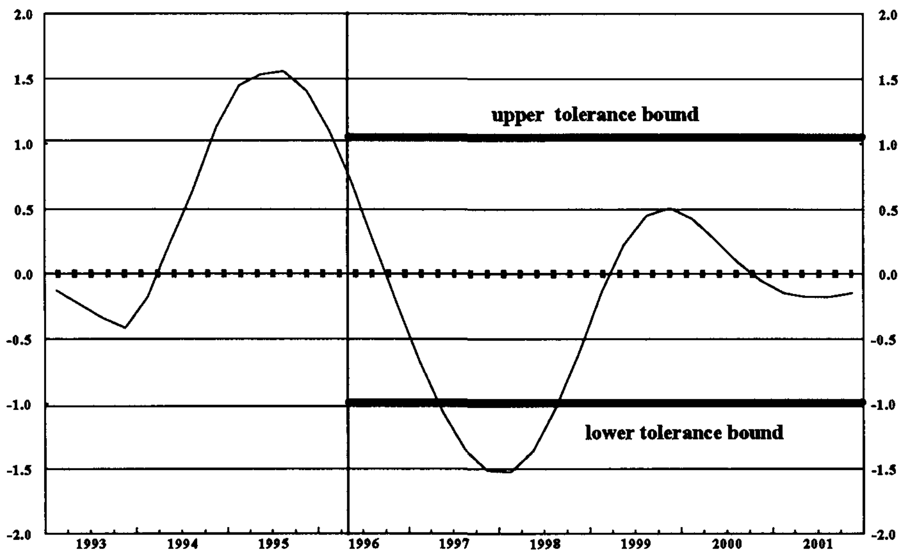


Figure 17D: One-year inflation expectation error

Political credibility shock case (%p.)



data suggests a period of about three years, beginning 1993, before expectations were in line with the inflation target.[15] Figure 17D shows the expectation error for a one-year inflation forecast associated with our political credibility shock scenario. Actual inflation now comes down within the tolerance interval and one-year inflation expectations are in line with the target about three quarters later, compared to the base case.

The corresponding graphs for expected inflation two years ahead are shown in Figures 18A-D, with the same general pattern of initially overestimating and later underestimating actual inflation evident from Figures 18A-B. Not surprisingly, inflation expectations for a two-year horizon show larger deviations from the target rate and settles down within the tolerance interval about a year later compared to one-year inflation expectations.

Actual inflation vs. two-year inflation expectations

Figure 18A: Base case (%)

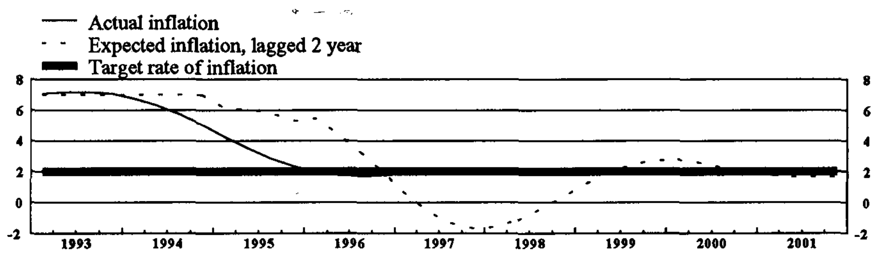


Figure 18B: Political credibility shock (%)

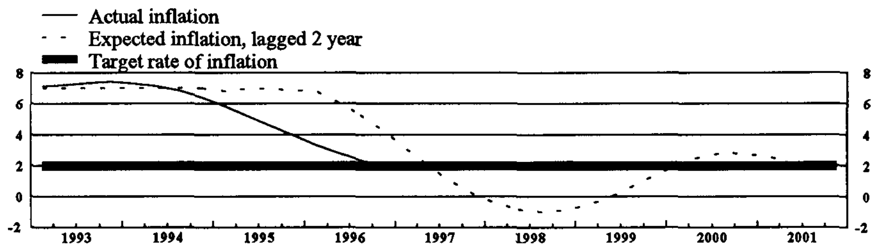


Figure 18C shows the expectation error for a two-year inflation forecast associated with our base case scenario. Compared to one-year inflation expectations it takes about one year longer for two-year expectations to be consistent with the inflation target. As with one-year expectations, political credibility shocks delay the settling down of expectations by slightly less than a year, as seen from Figure 18D.

15 See e.g. the March 1997 issue of the Inflation Report by Sveriges Riksbank.

Figure 18C: Two-year inflation expectation error

Base case (%p.)

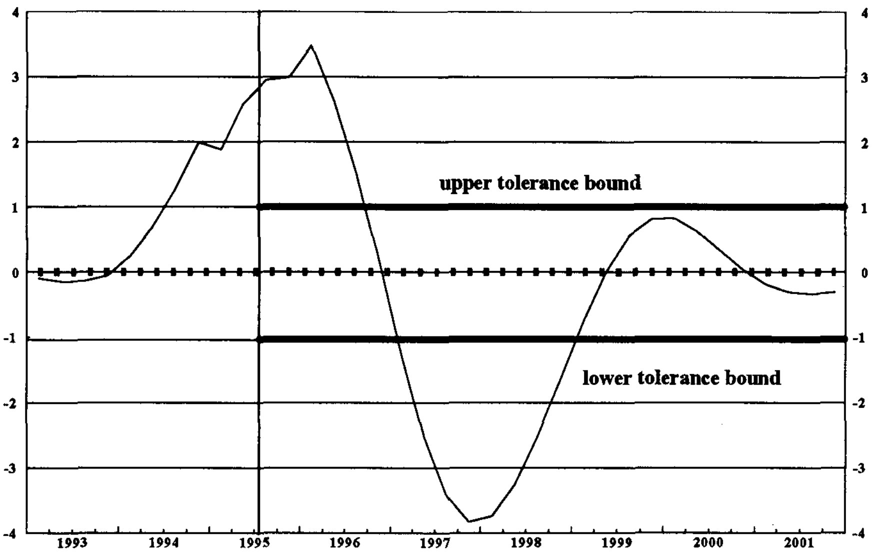
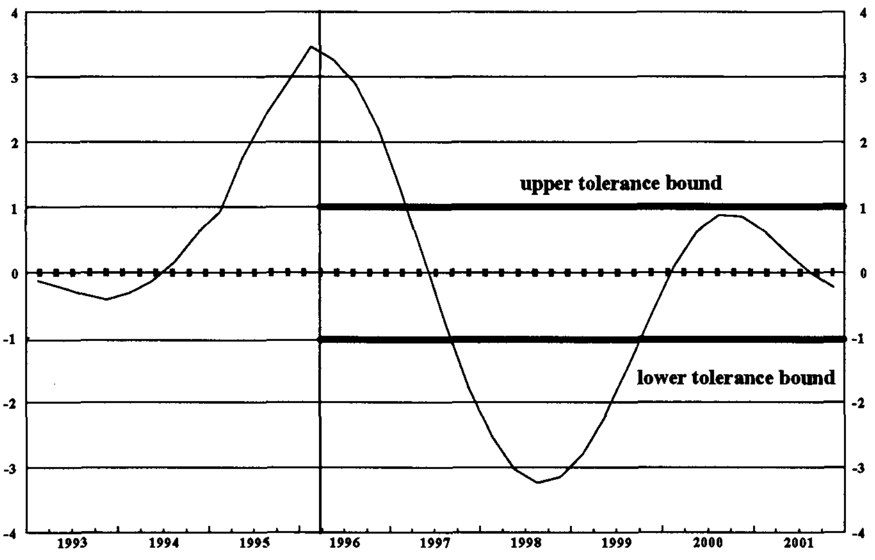


Figure 18D: Two-year inflation expectation error

Political credibility shock case (%p.)



Conclusions

This paper sets out to study the effects of a lack of credibility during the process of establishing a new monetary regime in Sweden in the early nineties. We distinguished between operational or endogenous credibility effects on the one hand, and political or exogenous credibility on the other.

The former is closely associated with how expectations about future inflation gradually adjust in the light of the actual inflation performance of the monetary authority in response to a downward shift of the inflation target. This mechanism is treated endogenously in our model. Next, we introduced political credibility shocks in the form of exogenous disturbances to the nominal exchange rate and to the long-term interest rate, such that there was an initial depreciation of the exchange rate and a widening of the long-term interest rate differential vis-a-vis the rest of the world. Political credibility shocks of this kind could go a long way in explaining the observed positive correlation between exchange rates and interest rates in Sweden, especially in the beginning of the nineties.

The main conclusion is that, in addition to the costs already associated with a transition to a low inflation regime, these political credibility shocks have substantial real effects. Our simulations suggest that these additional costs, measured as cumulated output losses, are in the neighbourhood of 2% of GDP. It should be emphasised that this estimate is probably biased upwards, since we have assumed that the regime shift is completely unannounced and initiated from a situation characterised by full employment of all resources. A more realistic simulation, starting from a disequilibrium, would probably reduce this figure.

Appendix

The one and two-year expected inflation rates are defined respectively as follows:

$$E_{t+i}(\pi_{t+i+4}) = \left[\frac{E_{t+i}(P_{t+i+4}) - E_{t+i-4}(P_{t+i})}{E_{t+i-4}(P_{t+i})}\right] \quad i = 0,\ldots,100$$

$$E_{t+i}(\pi_{t+i+8}) = \left[\frac{E_{t+i}(P_{t+i+8}) - E_{t+i-4}(P_{t+i+4})}{E_{t+i-4}(P_{t+i+4})}\right] \quad i = 0,\ldots,100$$

where $i = 0$ represents the current period. Thus, the inflation rates expected one and two years into the future, measured from the current period, are given by the following:

$$E_t(\pi_{t+4}) = \left[\frac{E_t(P_{t+4}) - E_{t-4}(P_t)}{E_{t-4}(P_t)}\right]$$

$$E_t(\pi_{t+8}) = \left[\frac{E_t(P_{t+8}) - E_{t-4}(P_{t+4})}{E_{t-4}(P_{t+4})}\right]$$

Likewise, the inflation rates expected one and two years into the future, measured from the next period ($i = 1$), are given by the following:

$$E_{t+1}(\pi_{t+5}) = \left[\frac{E_{t+1}(P_{t+5}) - E_{t-3}(P_{t+1})}{E_{t-3}(P_{t+1})} \right]$$

$$E_{t+1}(\pi_{t+9}) = \left[\frac{E_{t+1}(P_{t+9}) - E_{t-3}(P_{t+5})}{E_{t-3}(P_{t+5})} \right]$$

$$DLPEAQ = E_{t-1}(\pi_t) = \chi \left[\frac{E_{t-1}(LP_{t+20}) - E_{t-1}(LP_t)}{20} \right] + (1-\chi)\pi_t^{target}$$

where $\pi_t^{target} = \sum_{i=3}^{12} PTAR(t-i)/40$

$$DLPEA = 4 \cdot DLPEAQ$$

$$LP\_DA\_DIFF = DLPEA - DLPROWEA$$

References

Andersson, K. and C. Berg (1995): "The Inflation Target in Sweden", in A. Haldane (ed.), *Inflation Targeting*, Bank of England.

Bertola, G., and L. E. O. Svensson (1993): "Stochastic Devaluation Risk and the Empirical Fit of Target Zones Models". *Review of Economic Studies*, No. 60, pp. 689-712.

Black, R., D. Laxton, D. Rose and R. Tetlow (1994): "The Bank of Canada's New Quarterly Projection Model, Part 1, The Model of the Steady State: SSQPM". *Technical Report*, No.72, Bank of Canada.

Coletti, D., B. Hunt, D. Rose and R. Tetlow (1996): "The Bank of Canada's New Quarterly Projection Model, Part 3, The Dynamic Model: QPM". *Technical Report*, No.75, Bank of Canada.

Dillén, H. (1994): "A Model of the Term Structure of Interest Rates in an Open Economy with Regime Shifts". *Working paper*, No. 21, Department of Economics, Uppsala University. Forthcoming in *the Journal of International Monetary and Finance.*

Dillén, H. (1996): "Regime Shift Premia in the Swedish Term Structure: Theory and Evidence". *Working paper*, No. 28, Economics Department, Sveriges Riksbank.

Dillén, H. and E. Hopkins (1997): "Regime Shift Premia, Forward Interest Rates and Inflation Expectations". Unpublished manuscript, Economics Department, Sveriges Riksbank.

Dillén, H. and H. Lindberg (1997): "Interest and Exchange Rates in an Economy with Regime Shifts and Sticky Prices". Unpublished manuscript, Economics Department, Sveriges Riksbank

Svensson, L.E.O (1995): "The Swedish Experience of an Inflation Target", in L. Leiderman and L.E.O. Svensson (eds.), *Inflation Targets*, CEPR, London.

Sveriges Riksbank (various issues): "Inflation Report", Sveriges Riksbank, Stockholm.

Comments on: "Disinflation and credibility effects: the Swedish case" by Hans Dillén, Tohmas Karlsson and Jonny Nilsson

---

by Stefan Gerlach

The purpose of this paper is to study the macroeconomic effects of "credibility" shocks, by simulating the Riksbank's macroeconomic model which is very closely related to the Bank of Canada's QPM model. While this topic is of broad interest for many countries, it is particularly relevant in the Swedish case, where the credibility of monetary policy was limited after the floating of the krona in 1992, and where expectations as to whether the Riksbank will reach its inflation target have shifted over time. Of course, the topic is also of interest from a modelling perspective; shifts in investors' expectations of the likely future course of policy are likely to be common, and it is therefore important to know how to study them.

In the first section of the paper the authors briefly review the credibility of Swedish monetary policy since the floating of the krona in 1992, and argue that this period can be divided into two parts. Defining credibility shocks as disturbances that give rise to a positive comovement of the krona/DM exchange rate and the long-term interest rate differential with Germany, the authors show that between the end of 1992 and the summer of 1994, credibility shocks were relatively modest. Since then, however, there have been a number of shocks, and credibility has risen considerably.

I have two minor comments with this exercise. First, one should recall that monetary policy in Germany has been relaxed gradually since 1992. There are good reasons to believe that this helped underpin credibility of monetary policy in Sweden. Thus, it is not clear that the credibility shocks are all domestic, as implicitly assumed by the authors.

Second, the authors argue that under floating rates, a lack of credibility is likely to show up at the far end of the yield curve (in contrast to under fixed rates, when it shows up in very short-term interest rates). In light of this, and the fact that the Bank of Sweden does estimate zero coupon yield curves for several countries, it would be interesting to see what Figure 2 would look like if instantaneous forward rates, say ten years from now, were plotted instead of long bond yields.

In Section 2 of the paper the authors provide a clear overview of the model used for the simulations. The authors indicate that expectations are modelled as having two components: a forward-looking component, which is said to be *model consistent*, and a backward-looking component, which is *adaptive*. The authors seem too apologetic about the use of backward looking expectations. It is well-known that when the economy is disturbed by a mixture of permanent and transitory shocks, it is rational to form expectations about the realisation of the permanent shocks adaptively.

This is of course precisely the scenario considered in the paper: there is a permanent shift in the central bank target rate of inflation, and this policy change lacks credibility precisely because the public doesn't now yet whether the shift is permanent or transitory. The notion that expectations are partially backward-looking is therefore quite natural in this context.

In the simulation the authors consider a permanent reduction of inflation from seven to two percent, and show that the economy adjusts gradually to the policy change. It was not clear to me if the gradual response occurred solely because expectations were partially backward looking, or if there were other causes of the lagged response. Some clarification of this issue would be desirable.

Next follows an interesting discussion of how to model credibility shocks. The authors do so by shocking the nominal exchange rate and the nominal long-term interest rate. Since these variables are endogenous, this exercise strikes me as bit strange; I would have preferred if the authors left these variables unconstrained, and used them to see if they respond to the credibility shocks as one would expect on the basis of theory.

Finally, a few words on the simulation results. The key finding is that a lack of credibility makes disinflation costlier than otherwise. Of course, this finding is neither new nor surprising. What is new, however, are the quantitative estimates of these effects: the results suggest that lack of credibility makes the recession induced by the disinflation deeper by about 1.5% of GDP, and 1-2 years longer. It would be interesting to see how sensitive these results are to the alternative assumptions regarding the degree to which expectations are backward-looking. It would also be interesting to see whether it is optimal for the central bank to disinflate quickly or slowly.

Wealth and the demand for money in the European Union

Martin M. G. Fase and Carlo C. A. Winder[1]

Introduction

Under the Maastricht Treaty, the member states of the European Union have agreed to a single European currency and in 1999 at the latest on a common monetary policy aimed at price stability. Since this historical moment is rapidly approaching and because of the far-reaching nature of this commitment, the monetary authorities will have to decide soon which monetary policy strategy should be adopted. Price stability has been endorsed by all member states as the ultimate policy objective. The discussion focuses on the issue whether intermediate monetary targeting or direct inflation targeting should be chosen. In both monetary strategies monitoring monetary aggregates is important but will undoubtedly play a more prominent role in intermediate monetary targeting. The choice of a specific strategy will be crucial for the presentation of monetary policy measures and thus will affect directly the degree of transparency, accountability and credibility of the European central bank. Obviously, the choice between the two strategies will also be determined by an assessment of their effectiveness to achieve price stability. For a strategy of intermediate monetary targeting to be effective, a number of conditions have to be satisfied. The most important are the stability of the demand for money and the controllability of the money stock. Hence, it is not surprising that the demand for money has been again a major research topic in recent years (see e.g. Bekx and Tullio (1989), Kremers and Lane (1990), Monticelli and Strauss-Kahn (1992), Fase (1994, 1995) and Fase and Winder (1993, 1994)). The present study aims at contributing to the discussion by analysing the demand for money in the European Union as a whole, with a particular focus on the impact of financial wealth. In this respect, our analysis differs from earlier studies which have thus far neglected wealth as a determinant of the demand for money.

A stylised fact of the monetary developments in the European Union is the strong increase of notably the broad monetary aggregates M2 and M3, particularly since the beginning of the 1980s. This growth has exceeded to a considerable extent the growth of nominal output. According to the monetarist view on inflation, this monetary expansion should have led to an upward tendency in the inflation rates, but the 1980s have actually shown a sharp reduction of inflation. In order to assess the merits of an intermediate monetary targeting policy, it is highly desirable to have an economic explanation of this strong increase of the monetary aggregates. This study will argue that the development of financial wealth offers an explanation of this phenomenon. Taking account of wealth will lead to the conclusion that the actual monetary expansion has been fairly modest and that the strong increase of the broad monetary aggregates should be attributed to portfolio investment considerations rather than to an expansionary monetary policy.

The paper is organised as follows. Section 1 briefly discusses the data and presents some tentative empirical evidence concerning the impact of wealth on the demand for money. Section 2 discusses the theoretical framework and the empirical results for the European Union as a whole. Section 3 is devoted to an assessment of the findings. First, we will present evidence on the stability of the demand for money. Next, we will go further into the role of wealth and, finally, we will briefly discuss the developments of Divisia aggregates for the European Union. In principle, these aggregates offer an alternative way of taking into account the impact of portfolio investment considerations on

---

[1] The authors are economists in the Econometric Research and Special Studies Department of De Nederlandsche Bank NV, and, in addition, M. M. G. Fase is a member of the Faculty of Economics and Econometrics at the University of Amsterdam.

the demand for money. The final section concludes. Specific statistical issues and additional estimation results are set out in three appendices.

1. Construction of data and preliminary analysis

For the present study the database used in Fase and Winder (1993,1994), containing data for all EU-countries, before the latest enlargement, for the monetary aggregates M1, M2 and M3, gross domestic product in current and 1985-prices and short and long-term interest rates, have been extended with information for Austria, Finland and Sweden. Hence, data for all present members of the European Union with the exception of Luxembourg have been used to construct EU-wide variables. The sample period is 1971Q1-1995Q4. In order to obtain a comprehensive picture of the monetary developments we did not restrict the analysis to one monetary aggregate but considered a broad spectrum; i.e. three money concepts: M1, M2 and M3. To investigate the impact of wealth on the demand for money we also constructed wealth data for all individual countries. Appendix A contains an extensive discussion of the procedure to construct these data. Following the seminal study of Brainard and Tobin (1968) we considered net financial wealth of the non-monetary private sector as the relevant wealth variable. This is defined as the difference between total assets – the sum of M1, quasi-money, claims on the banking sector and on the government and net foreign assets – and the claims of the monetary authorities and the banking sector on the private sector. A more general wealth concept, which, besides financial wealth, also includes non-financial components like the capital stock, houses and perhaps durable consumption goods, has not been considered, because data for these debt instruments are not readily available. One may argue that since portfolio investment considerations dominate money demand behaviour, it is justified to focus on financial wealth. However, even if a broader wealth concept would be more appropriate, the results in the present study remain valid if financial wealth is a good proxy for total wealth.

For the construction of EU-wide data, several options are available. In order to measure monetary aggregates, nominal income and financial wealth of the individual countries in a common currency we used exchange rate and purchasing power parities against the DM. For both conversion methods one may use actual, i.e. current, values or values in a base period, resulting in a total of four different options for conversion. After expressing the national variables in a common currency, the EU-wide variables are calculated as the sum of these national variables. The figures of the countries' national product in 1985 prices have been converted into 1985 DM prices using the exchange rates and purchasing power parities against the DM in 1985. EU-output in 1985 prices has been obtained by summing these transformed data on national products in 1985 prices. The price index for EU has been calculated as the ratio between EU-output in current and constant prices.

The short and long-term interest rates for the EU have been calculated as a weighted average of the corresponding interest rates in the individual countries. As weights we considered the shares of the national products in the individual countries in the aggregated EU-output. Since EU-output has been calculated in four different ways, viz. using current and base period exchange rates and purchasing power parities, we have also four different weighting schemes for the short and long-term interest rates.

The various possibilities for constructing EU-wide variables lead to four alternative series for monetary aggregates M1, M2 and M3, nominal output and wealth. On the basis of theoretical arguments, it is not possible to make a definite choice. The selection of the most appropriate construction method is therefore mainly an empirical matter. The results reported in the main text concern the data constructed with 1985 DM exchange rates as conversion factor.[2]

---

[2] However, if alternative calculation methods lead to sharply differing results, the question may arise whether an analysis at EU-level is meaningful at all. To address this issue we performed a preliminary analysis of the EU-wide

The remainder of this section is devoted to a preliminary analysis of the impact of wealth on the demand for money. Particularly, we address the issue whether the development over time of wealth offers an explanation for the monetary expansion experienced in the European Union over the last twenty years. To throw some light on this issue the development of the wealth-income ratio and the liquidity ratio, i.e. the money stock as a percentage of output in nominal terms, are depicted in Figure 1. The shaded bars in the panels denote the periods in which for more than two subsequent quarters the short-term interest rate was higher than the long-term interest rate; i.e. the yield curve had a decreasing – or inverse – slope.

Figure 1 shows that the wealth-income ratio has increased strongly in the first half of the 1980s. For the period 1987-93, this ratio has, as a matter of fact, been constant but from 1993 onwards has shown a steady increase. In contrast, the wealth-income ratio has been relatively constant during the 1970s. Obviously, wealth has evolved far from proportional to nominal income, thus causing serious doubts about the often-made assumption in the relevant literature that income can be used as a proxy for wealth. The liquidity ratios in Figure 1 illustrate the monetary expansion in the European Union, which has occurred especially for the broader monetary aggregates M2 and M3. A salient feature of Figure 1 is the diverse developments of the liquidity ratios. The M1-ratio has been rather volatile, though within relatively narrow margins, but has been lower in the 1980s and 1990s than in the 1970s. In contrast, the M2 and M3-ratios have increased steadily over the period considered, though at the end of the sample period a stabilisation or even small decrease occurs. The rise of the M2-ratio was relatively modest during the 1970s and accelerated from the beginning of the 1980s onwards. The M3-ratio, on the other hand, shows a more steady rise during the sample period.[3]

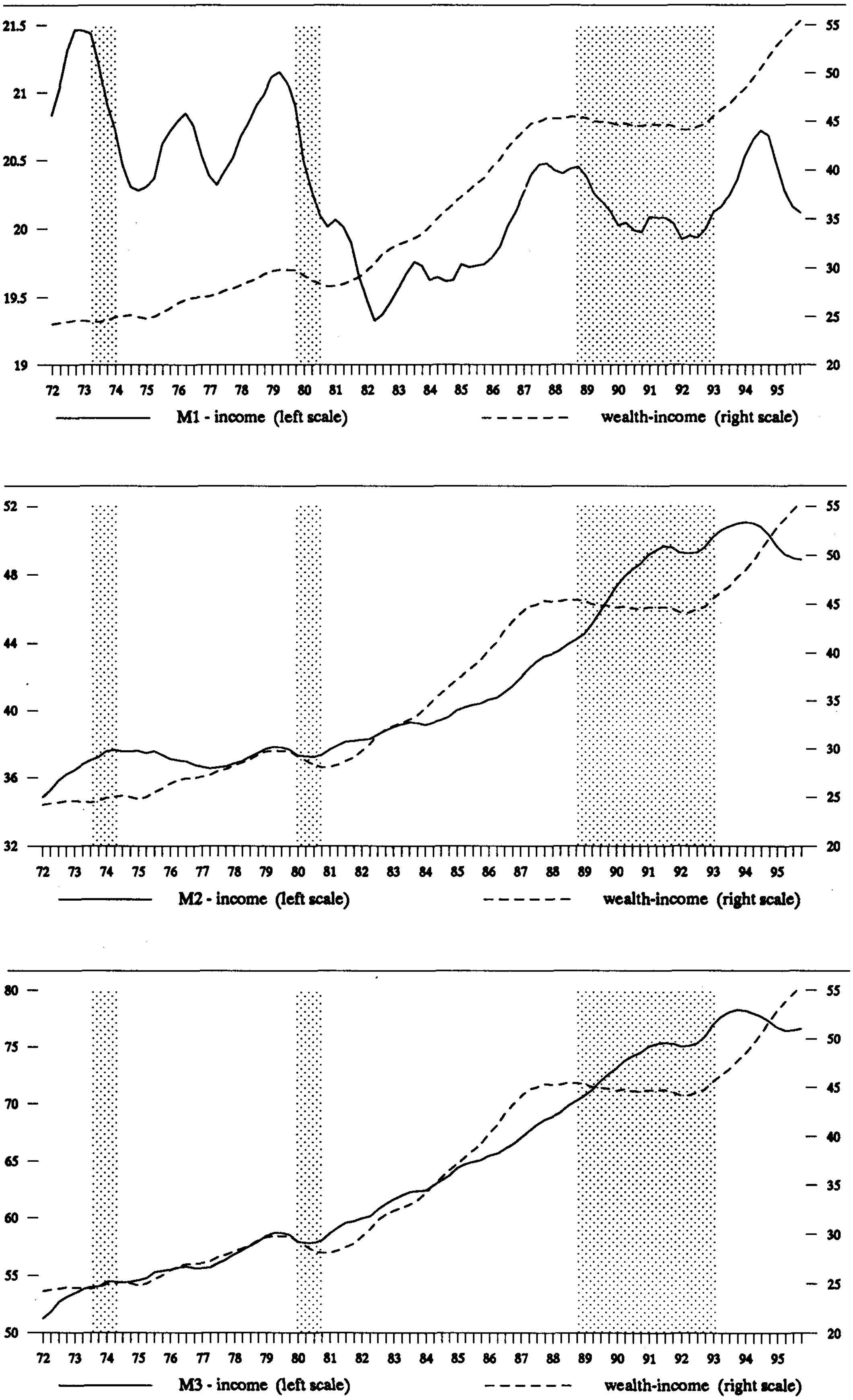
A feature of Figure 1 that catches the eye is the similar development of the M2 and M3-ratios on the one hand and the wealth-income ratio on the other hand. The development of the M1-ratio is much less in line with that of the wealth-income ratio. Obviously, the decrease of the M1-ratio in the beginning of the 1980s is not related to the development of wealth. The resulting pattern for the broader monetary aggregates, however, supports the hypothesis that the relatively strong increase of wealth offers an explanation for the rising M2 and M3-ratios. In particular, during the period 1972-87 the liquidity ratios for M2 and M3 have evolved parallel to the wealth-income ratio. In the period 1987-1992 the wealth-income ratio stabilised, whereas the M2 and M3-ratio show a continuing increase. It is likely that this further rise is related to interest rate developments. In all EU-countries the short-term interest rate increased sharply in this period leading to inverse yield curves. The resulting shift from capital market investments to near money led to high growth rates for the broad monetary aggregates, and hence to a further increase of the M2 and M3-ratios. Periods with an inverted yield curve are denoted in Figure 1 by the shaded bars, showing that for two other subperiods the short-term interest rate exceeded the long-term rate. The effect of the interest rate developments on the liquidity ratios in these periods is much less clear cut, perhaps because their length was considerable shorter than for the period around 1990. The adverse development of the liquidity ratios for the broad aggregates and the wealth-income ratio during the period 1987-93 seems, however, to be related to the development of the interest rates. This illustrates that to assess the monetary

---

variables constructed with 1985 DM purchasing power parities, current DM exchange rates and current DM purchasing power parities as conversion methods. The results are presented in Appendix A, showing that the stylised facts to be discussed in this section, are shared by the alternative conversion methods. The results in Appendix A warrant, therefore, the conclusion that specific choice of aggregation method, though having influence on details, does not affect the overall pattern in such a way that it is prohibitive for an analysis at EU-level.

3 The development of the EU-aggregates conceals substantial differences between the individual countries. As documented in Fase and Winder (1993), monetary expansion was especially strong in the northern EU-countries, notably Denmark, the Netherlands and the United Kingdom, whereas the southern countries offer a relatively stable picture. Differences between the demand for money relationships for the individual countries give rise to an aggregation bias in the estimates of EU-wide demand for money equations. Aggregation may on the other hand reduce the specification bias. As a result of the trade-off between aggregation and specification bias, EU-wide estimates may perform better than single-country ones (see Kremers and Lane (1990)).

Figure 1
Liquidity ratios and wealth-income ratio



developments it is necessary to take into account other determinants of the demand for money than income and wealth. The theory of the demand for money offers an analytical framework for such a more comprehensive analysis. The results of this analysis will be discussed in the next section.

2. Theoretical framework and empirical results

One of the reasons for the long history of research on the demand for money in empirical econometrics is the availability of a reasonably well-developed theory. This is supported by the historiography of views on the demand for money. The history of the neo-classical theory of the demand for money starts with the quantity theory of money proposed by e.g. David Hume and Irving Fisher, usually summarised by the latter's equation of exchange which opposes the flow of money and the flow of goods and services. This implies that the volume of current transactions determines the demand for money. The ensuing Cambridge variant, known as the cash balance theory with the cash ratio as the crucial parameter depending on interest rates, paid explicit attention to money as a store of value and thus wealth. A major step forward in the development of the ideas on the demand for money was Keynes' liquidity preference theory with the famous three motives for holding money, resulting in income, interest rate and wealth as determinants of demand for money. Friedman's reformulation of Fisher's quantity theory considered the demand for money as a special case of the demand for goods and services. In this approach, the demand for money is just one equation from a complete set of demand equations. Consequently, Friedman's theory was closely related to the portfolio approach of e.g. Tobin which naturally included wealth. Therefore, Friedman's and Tobin's approach explicitly takes wealth into account to explain money holding behaviour. In this view, money is the most liquid financial asset in the whole spectrum of assets.

Following Friedman's theoretical framework, the postulated long-run demand for money relationship reads as:

$$M = P^{\alpha_1} y^{\alpha_2} w^{\alpha_3} \exp(\alpha_4 rs + \alpha_5 rl + \alpha_6 \dot{p}) \tag{1}$$

with M the desired nominal money stock in a steady state situation, P the price level, y real income, w real wealth and $\dot{p}$ the inflation rate while rs and rl represent the short-term and long-term interest rates. The price elasticity equals α_1, while α_2 and α_3 represent the income and wealth elasticities; α_4, α_5 and α_6 measure the relative change of money demand as a result of a one percentage point increase of the short-term interest rate, the long-term interest rate and the inflation rate respectively. These parameters are therefore the semi-elasticities with respect to the short and long-term interest rates and inflation. The often made assumption of price-homogeneity, i.e. α_1=1, is not imposed as an a priori restriction, but considered as an empirically testable hypothesis. If this restriction is valid, the steady-state relationship of equation (1) relates real money balances to the explanatory variables, with y reflecting the impact of the transactions volume and wealth the scale variable reflecting portfolio investment considerations. An alternative interpretation of this latter determinant is that it is a proxy for the influence of the money demand resulting from financial transactions (see also Friedman (1971)). These transactions are not accounted for appropriately by the income variable. The interest rates and the inflation rate represent the substitution processes between financial and physical assets. The short-term interest rate is the appropriate yield on near money. For the narrow monetary concept M1, α_4 should therefore be negative, while for the broader aggregates M2 and M3 a positive sign is plausible on theoretical grounds. The semi-elasticities α_5 and α_6 should be negative for M1, M2 and M3. Including both the short and long-term interest rate in the analysis takes into account also the influence of the slope of the yield curve on the demand for money.

Equation (1) is the maintained hypothesis and describes the steady state demand for money relationship. In case of deviations from this long-run equilibrium, adjustment processes occur to restore equilibrium. These short-run adjustment processes are modelled by means of lagged values of the variables in equation (1), leading to an equation in which the short and long-run properties of the demand for money are described jointly. Following Friedman (1971) and earlier studies carried out at the Bank, a cyclical indicator, *conj*, has also been included to describe these short-run adjustment processes (see e.g. Fase and Kuné (1974) and Den Butter and Fase (1981)). This leads to the following dynamic equation:

$$\alpha(L)\ln M = \beta_1(L)\ln P + \beta_2(L)\ln y + \beta_3(L)\ln w + \beta_4(L)rs + \beta_5(L)rl + \beta_6(L)\dot{p} + \beta_7(L)conj \qquad (2)$$

with L the lag operator and $\alpha(L)$ and $\beta_i(L)$, i=1,..,7, lag polynomials. Fase and Winder (1993) discuss the identification problem concerning the parameters of the lag polynomials $\beta_1(L)$ and $\beta_6(L)$ (see also Filosa (1995)). In line with their analysis a number of a priori restrictions have been imposed, leading to:

$$\alpha(L)\ln(M/P) = (\beta_{10} - 1)\ln P + \beta_2(L)\ln y + \beta_3(L)\ln w + \beta_4(L)rs + \beta_5(L)rl + \beta_6(L)\dot{p} + \beta_7(L)conj \qquad (3)$$

with β_{10} the constant of the lag polynomial $\beta_1(L)$. Reformulating (3) yields the estimation equation considered in the empirical analysis:

$$\begin{aligned}\overline{\alpha}(L)\Delta\ln(M/P) = {} & (\beta_{10} - 1)\ln P - \alpha(\ln(M/P)_{-1} - \ln y_{-1}) + \beta_3(\ln w_{-1} - \ln y_{-1}) + (\beta_2 + \beta_3 - \alpha)\ln y_{-1} \\ & + \beta_4 rs + \beta_5 rl + \beta_6\dot{p} + \overline{\beta}_2(L)\Delta\ln y + \overline{\beta}_3(L)\Delta\ln w \\ & + \overline{\beta}_4(L)\Delta rs + \overline{\beta}_5(L)\Delta rl + \overline{\beta}_6(L)\Delta\dot{p} + \beta_7(L)conj \qquad (4)\end{aligned}$$

with β_i, i=2,...,6, equal to the sum of the coefficients of the lag polynomial $\beta_i(L)$ and α the sum of the coefficients of $\alpha(L)$. From equation (4) the long-run (semi)-elasticities can easily be determined. The semi-elasticities with respect to the short and long-term interest rates and the inflation rate are equal to β_4/α, β_5/α and β_6/α respectively. The price elasticity equals $(\beta_{10} - 1 + \alpha)/\alpha$, whereas the equilibrium income and wealth elasticities are equal to β_2/α and β_3/α. With equation (4) it is relatively simple to test restrictions with respect to the long-run price, income and wealth elasticities. The hypothesis that the price elasticity equals 1 can be tested by assessing the significance of the estimate of the coefficient of ln P. The hypothesis that the wealth elasticity equals 0 corresponds to the restriction that the coefficient of $(\ln w_{-1} - \ln y_{-1})$ is zero. The hypothesis that the income and wealth elasticities add to 1 – a restriction which is often imposed but for which, incidentally, no theoretical argument is available – corresponds to the restriction that the coefficient of $\ln y_{-1}$ is equal to 0. Obviously, these restrictions can easily be tested by means of the t-statistics of the estimated parameters.

The demand for money in the European Union as a whole and particularly the impact of wealth, have been analysed using equation (4). In our discussion of the empirical results we will focus on the long-run properties implied by the estimated demand for money relationships. In order to gain insight into the properties of the demand for money, a discussion of the steady state is obviously most revealing. To this end, the estimated parameters of the steady state demand for money relationships for the various monetary aggregates are summarised in Table 1. The detailed estimation results for M1, M2 and M3 are reported in Appendix B. The results of the statistical evaluation of the estimated equations are given in Appendix C. These show that from a statistical point of view the specification of the money demand equations is appropriate.

As Table 1 shows, wealth has a significant impact on the demand for M2 and M3. For M1 no influence of wealth was found, a result that was already suggested before by Figure 1. The income elasticity of the demand for M1 is higher than for M2 and M3, indicating that M1-balances are mainly held for transaction purposes. For M2 and M3, quite arbitrarily, the restriction has been imposed that the income and wealth elasticities add to 1. This restriction constitutes a testable hypothesis and the results reported in Appendix C show that this restrictive hypothesis cannot be rejected. The finding that the income and wealth elasticities add to 1 corresponds to the results reported for various individual European countries. Examples are the Bundesbank (1995) for Germany, Hall et al. (1989) for the United Kingdom, and Fase and Winder (1996) for Belgium and the Netherlands (see also Sterken (1992)). Fase and Winder found that the wealth elasticity of the demand for M2 was slightly higher than for M3 in both countries. In contrast, the results in Table 1 show that for the European Union as a whole the wealth elasticity of M3 is somewhat higher. Furthermore, for all monetary aggregates considered in this paper, the price homogeneity restriction has been tested and cannot be rejected (see Appendix C).

Table 1
Estimates of parameters of long-run money demand relationship

Monetary aggregate	α_1	α_2	α_3	α_4	α_5	α_6
M1	1	0.8619	0	-0.0086	-0.0052	-0.0052
		(15.14)		(2.89)	(1.74)	(1.74)
M2	1	0.7479	0.2521	0.0126	-0.0195	-0.0195
		(11.25)	(3.79)	(1.99)	(3.44)	(3.44)
M3	1	0.6590	0.3410	0.0107	-0.0133	-0.0133
		(12.19)	(6.31)	(2.28)	(2.91)	(2.91)

Note: t-values are given in parentheses.

With respect to the semi-elasticities of the short and long-term interest rate and the inflation rate, all the estimates obtained have the theoretically plausible sign and are, with the exception of the semi-elasticities of M2 with respect to the long-term interest rate and inflation, significantly different from 0. For M1, M2 and M3 the results of the econometric specification analysis (see Appendix C) allow us to impose the restriction that the inflation and long-term interest rate semi-elasticities are equal. As Table 1 shows, the demand for M2 is more sensitive to changes in the interest rates and the inflation rate than the demand for M1 and M3. In fact, M1-balances have relatively low interest rates and inflation semi-elasticities, a finding that corresponds with the notion that these balances are mainly held for transaction purposes.

An alternative way to assess the estimation results is to consider the properties of near money. From the elasticities for M1, M2 and M3 one can derive demand elasticities for M2–M1, M3-M1 and M3–M2 and these elasticities must, of course, be economically plausible.[4] This offers the opportunity to assess the mutual consistency of the results obtained for M1, M2 and M3. Therefore, the elasticities of the various near-money concepts are presented in Table 2 together with the elasticities for M1, M2 and M3. The price, income and wealth elasticities follow from Table 1, whereas the elasticities with respect to the short and long-term interest rates and the inflation rate have been calculated by multiplying the semi-elasticities given in Table 1 by the sample means of the corresponding variables.

As Table 2 shows, the elasticities of the various near-money concepts have an economically plausible sign and a reasonable size, implying that the results for M1, M2 and M3 are mutually consistent. The price elasticities of the various near monies are equal to 1, reflecting the similar properties of M1, M2 and M3. The income elasticities of the various near-money concepts are smaller than for M1, with the demand for short-term time deposits and foreign currency deposits,

4 It can easily be verified that dln M2 = w dln M1 + (1-w) dln (M2-M1) with w = M1/M2. Similar relationships hold for the other assets. In the calculation of the elasticities of the various near-money concepts, the sample means of w have been used.

i.e. M2-M1, having a higher income elasticity than the demand for short-term saving deposits, i.e. M3-M2. A striking feature of Table 2 is the similar values of 0.50 for the wealth elasticities of the demand for the various near monies. The sum of the income and wealth elasticities equals 1 for M3-M2. This result reflects the corresponding property for the demand for M2 and M3. For the other two near-money concepts the sum of the income and wealth elasticities is slightly larger than 1, as the income elasticity of M1 is smaller than 1 while the income and wealth elasticities for M2 and M3 add to 1.

Table 2
Long run elasticities

Monetary aggregate	Price	Income	Wealth	Short-term interest rate	Long-term interest rate	Inflation
M1	1	0.862	0	-0.082	-0.053	-0.035
M2	1	0.748	0.252	0.120	-0.199	-0.129
M3	1	0.659	0.341	0.102	-0.136	-0.088
M2–M1	1	0.636	0.501	0.319	-0.343	-0.223
M3–M1	1	0.562	0.504	0.190	-0.175	-0.114
M3–M2	1	0.494	0.506	0.068	-0.019	-0.012

The general picture emerging from Table 2 for the interest rates and inflation elasticities of M1, M2 and M3 is similar to that found for the semi-elasticities discussed before. This is not surprising given the way these elasticities have been calculated. A notable feature of Table 2 is that the demand for M2–M1 is particularly sensitive to changes in the interest rates and the inflation rate, while the interest rates and inflation elasticities of the demand for M3–M2 are very small.

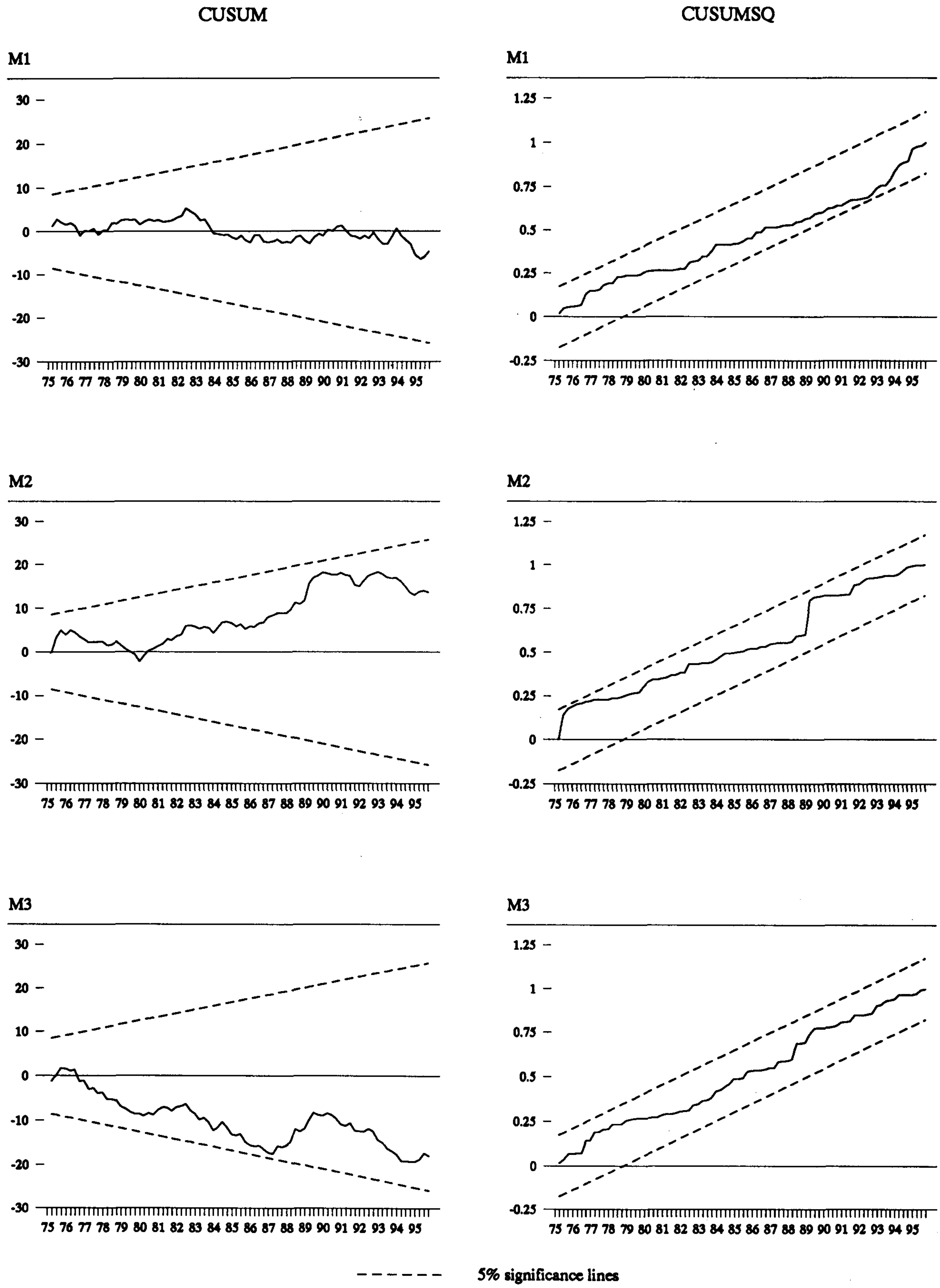
3. An assessment

In order to assess the results reported in the former sections we examine three issues. In the first sub-section we will present empirical evidence on the stability of the demand for money. The next sub-section will further analyse whether the development over time of wealth offers an explanation for the monetary expansion that has occurred in the European Union since the beginning of the 1980s. Sub-section 3.3 discusses the developments of Divisia aggregates for the European Union. It will be argued that, in principle, these aggregates offer an alternative way to take into account the impact of portfolio investment considerations on the demand for money.

3.1 Stability of the demand for money

An important assessment criterion, both from a policy and an economic viewpoint, is the degree of stability of the demand for money. Several methods are available to assess stability. Here we consider three methods. The first one concerns the residuals of the estimated equations. The standard error of the residuals of the estimated demand for money equations (Appendix B) are 0.67%, 0.80% and 0.58% for M1, M2 and M3 respectively. The size of these standard errors is rather low for all monetary aggregates considered, indicating that the overall fit is satisfactory. According to this measure of stability, the demand for M3 is more stable than the demand for M1 and M2, a finding that is in line with the results in Fase and Winder (1993) for the individual countries of the European Union. The residuals reflect the noise inherent in the estimated equations and may be useful in determining the width of the target zone if a policy of monetary targeting is pursued. However, this measure is less useful as a criterion to assess possible misspecifications of the chosen equations.

Figure 2
Recursive residuals



Therefore, we have also examined the stability over time of the estimated equations. This procedure involves re-estimating of the specifications for sub-periods and testing for parameter stability by means of Chow-tests. To this end, the equations have been re-estimated for the sample periods before and after the German unification in 1990 using 1990Q3 as a breakpoint. The hypothesis that the parameters in both subperiods are constant can be tested with an F-test, the F-statistic having an F(13,69) distribution for M1, M2 and M3. The values obtained are 2.27, 1.25 and 1.48 respectively. Only for M1 is the F-statistic significant at a 0.05 significance level, suggesting that the M1-equation is not stable. For M2 and M3 the hypothesis that the parameters are constant during both subperiods considered cannot be rejected. As a further test of stability we have also considered the constancy of the long-run parameters before and after German unification. From a policy viewpoint the long-run properties of the demand for money are more important than the short-run dynamics. The Chow-statistic in this case is F(3,79) distributed. This test yields the values 0.74, 0.65 and 0.20 for M1, M2 and M3 respectively, which are all insignificant at a 0.05 significance level. Hence the hypothesis that the long-run parameters are constant during the subperiods considered cannot be rejected. Moreover, the instability of the M1-equation found above is related to non-constancy of the parameters describing the short-run dynamics of the M1-equation.

As a final procedure to assess the stability of the demand for money we considered recursive estimation procedures, which are particularly useful in detecting structural changes. More specifically, we calculated the cumulative sum (CUSUM) and the cumulative sum of the squares (CUSUMSQ) of the recursive residuals.[5] Figure 2 depicts for M1, M2 and M3 the CUSUM and CUSUMSQ graphically, together with the 5% significance lines. For all monetary aggregates considered, neither the CUSUM nor CUSUMSQ cross the significance lines, suggesting that there is no evidence of parameter instability. This finding is encouraging and supports the earlier preliminary result that stability increases with the size of the economic geographical area considered using Divisia aggregates (see e.g. Fase (1995)). An explanation for this remarkable result found again in this paper is that the degree to which the supply of money is exogenous – strictly speaking a condition for the estimation of a single demand for money equation – is perhaps depending on the capacity to absorb shocks which certainly increases with the size of the economy.

3.2 Further assessment of the role of wealth

The empirical results in Section 2 revealed a substantial impact of wealth on the demand for M2 and M3, whereas for M1 no influence of wealth could be found. The tentative analysis in Section 1 showed a parallel development of the M2 and M3-ratios on the one hand and the wealth-income ratio on the other hand, suggesting that the increase in wealth offers an explanation for the rising liquidity ratios observed in the 1980s.

In the spirit of Fisher's quantity theory, liquidity ratios – or their inverse, i.e. the velocities of money – are often used as a simple indicator to assess monetary developments. It is well-known that these ratios are fairly rough measures because substitution processes resulting from changes of interest rates and inflation may cause large fluctuations of the money stock. Another reason to consider the liquidity ratios with some caution is that the assumption of a unit income elasticity of the demand for money may be violated. In this regard it may be recalled that for M1 the income elasticity is estimated to be 0.86. Moreover, the results in the present study show that wealth also exerts a major influence on the demand for money. Obviously, this influence should be taken into account when assessing the monetary developments.

The long-run demand for money relationship (1) implies:

$$M = P^{\alpha_1-\alpha_2-\alpha_3} Y^{\alpha_2} W^{\alpha_3} f(rs, rl, \dot{p})$$

---

[5] The CUSUM and CUSUMSQ have been calculated conditional on the dummy variable incorporated to account for the effect of the German unification (see Appendix B).

where Y and W are nominal income and wealth respectively, and $f(rs,rl,\dot{p}) = \exp(\alpha_4 rs + \alpha_5 rl + \alpha_6 \dot{p})$. Thus,

$$\frac{M}{Y} = P^{\alpha_1-\alpha_2-\alpha_3} Y^{\alpha_2-1} W^{\alpha_3} f(rs,rl,\dot{p})$$

implying that the liquidity ratio also depends on wealth. If the price elasticity α_1 equals 1 and the income and wealth elasticities sum to 1, i.e. $\alpha_2 + \alpha_3 = 1$,we have

$$\frac{M}{Y} = \left(\frac{W}{Y}\right)^{1-\alpha_2} f(rs,rl,\dot{p})$$

The relationship between the liquidity ratios and the wealth-income ratio has been considered in Figure 1. With the above restrictions concerning the price, income and wealth elasticities, we also have

$$M/\left(Y^{\alpha_2} W^{1-\alpha_2}\right) = f(rs,rl,\dot{p})$$

Since the interest rates and the inflation rate do not display trend-like behaviour, it may be expected that the same property holds for $\left(Y^{\alpha_2} W^{1-\alpha_2}\right)$. It may, therefore, be interesting to consider this ratio for the monetary aggregates. It goes without saying that this ratio is a purely descriptive statistic.[6]

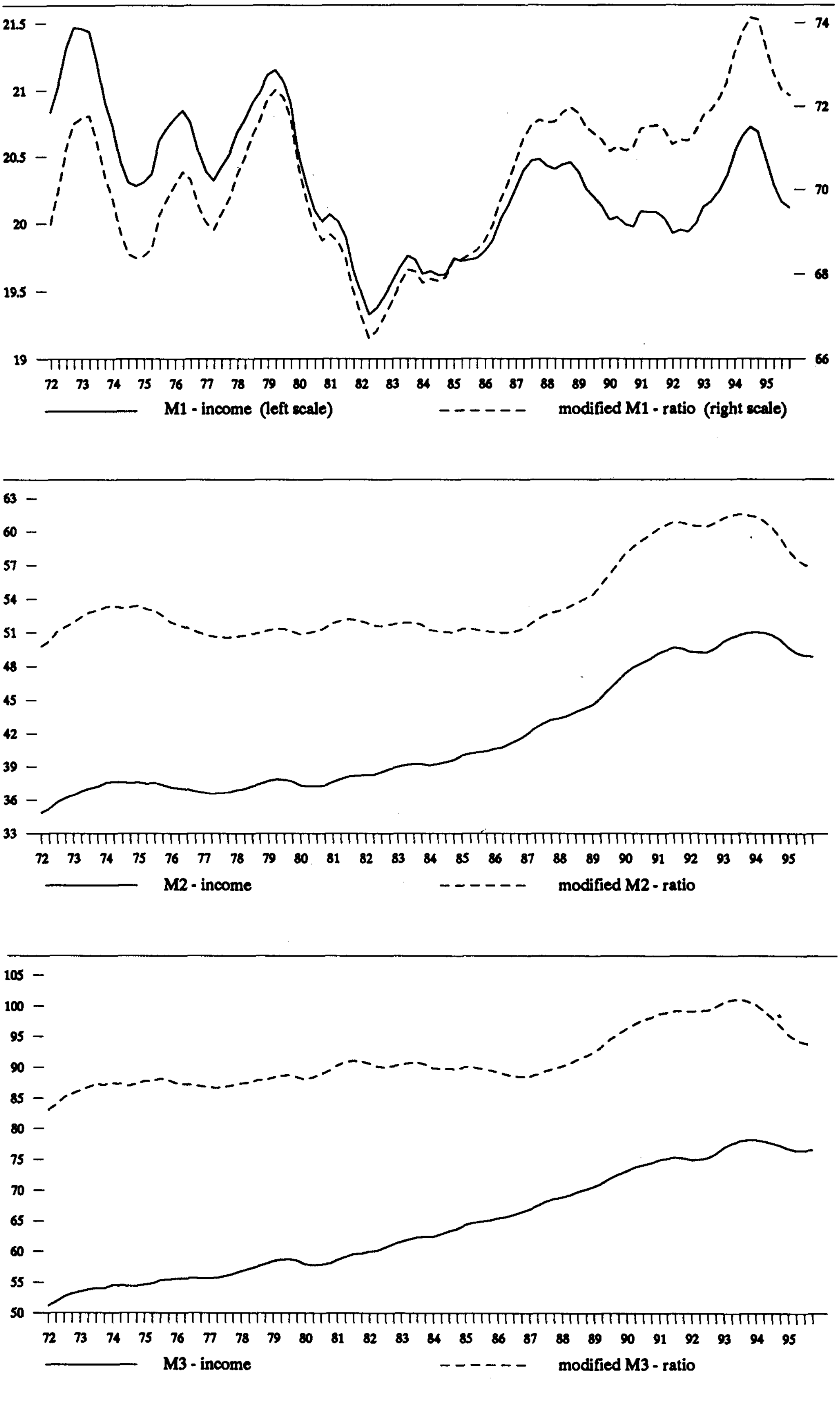
In Figure 3 it has been depicted together with the liquidity ratio for each of the money concepts considered, using the estimates of the long-run income elasticities given in Table 1. Since the wealth elasticity of M1 is 0, the ratio $M/\left(P^{1-\alpha_2} Y^{\alpha_2}\right)$, with α_2 being the long-run income elasticity, has been depicted. If the income and wealth elasticities add to 1, the modified ratio is independent of the units in which the money stock, income and wealth are expressed. This situation applies to M2 and M3. If the elasticities do not add to 1, as is the case for M1, the level of the modified ratio depends on the units in which the variables are expressed. Thus, it makes a difference whether M1 and income are expressed in millions or billions of German marks. It can easily be shown, however, that the development over time is not influenced by the choice of unit. It is only the level that is affected. Therefore, the M1-ratio, as well as the corresponding modified ratio, is depicted with a double scale. For M2 and M3 a single scale suffices.

As Figure 3 shows, the modified M2 and M3-ratios have been relatively constant over the 1970s and 1980s, but an increase at the end of the 1980s is clearly displayed, notably for the modified M2-ratio. The relatively strong increase of M2 and M3 is probably related to the influence of the inverted yield curve, which initiated shifts from capital market investments to liquid assets (see also Section 1). Like the M1-ratio, the modified M1-ratio has been rather volatile, albeit within relatively narrow margins. A salient feature of the modified M2 and M3-ratios in Figure 3 is that no upward tendency is present during the 1980s. This indicates that no excessive monetary expansion has occurred, a conclusion that is in line with the modest development of the inflation rates during the 1980s. Thus, the relatively stable developments of these modified ratios may be taken as a justification for the reserved reactions of the monetary authorities to the growth of the liquidity ratios.

---

6 If W is proportional to Y, $M/\left(Y^{\alpha_2} W^{1-\alpha_2}\right)$ will – apart from a constant – be equal to the liquidity ratio and therefore evolve parallel to the liquidity ratio.

Figure 3
Liquidity ratios and modified ratios



3.3 An alternative approach using Divisia monetary aggregates

In a study of Divisia aggregates for the European Union, Fase and Winder (1994) argued that the sharp increase of the liquidity ratios for the broad money concepts can be attributed to the increase of the volume of liquid assets held for portfolio investment purposes. When correcting M2 and M3 for these portfolio investment balances according to the methodology put forward by Divisia (1925, 1926), money growth relative to income growth has been modest during the 1980s, a conclusion that is in line with the results of the foregoing analysis. Here we will discuss the relationship between an analysis using Divisia aggregates and one which incorporates wealth as explanatory variable of the demand for money.

Divisia monetary aggregates are motivated by the notion that money can simultaneously serve various motives. It is a medium of exchange used for carrying out transactions, but it is also a store of value and, therefore, one way of holding wealth. Conventional monetary aggregates are constructed by summing the constituent liquid assets, assuming implicitly that these assets are perfect substitutes (see e.g. Barnett (1978)). In practice, this heroic assumption can only be maintained as an approximation. For instance, M1 is held primarily for transaction purposes, whereas for near monies investment considerations play a major role. Divisia aggregates correct for the differences in degrees of liquidity of the various liquid assets, aiming to provide a more appropriate measure of the volume of liquid assets held for transaction purposes. Obviously, the monetary authorities are especially interested in a monetary aggregate which is adjusted to eliminate the growth ensuing from investment considerations. Since Divisia aggregates aim to correct the conventional money concepts for portfolio investment changes, it can be argued that both an analysis with Divisia aggregates and one that includes wealth as a determinant of the demand for money address the same issue, viz. taking into account portfolio investment considerations. One approach includes wealth among the set of explanatory variables, while the alternative corrects conventional aggregates.

In order to investigate this issue further we have updated the results in Fase and Winder (1994) to take account of the enlargement of the EU with Austria, Finland and Sweden and the developments since 1992.[7] The liquidity ratios for the Divisia aggregates corresponding to M2 and M3 are depicted in Figures 4.1 and 4.2. As the Divisia aggregates represent index numbers, with the 1972Q1 level set at 100, the liquidity ratios for M1, M2 and M3 have been similarly scaled. The figures also depict the development of the modified M2 and M3-ratios discussed before, again after similar scaling.

The results in Figures 4.1 and 4.2 are very similar to those reported in Fase and Winder (1994). Both M2 and M3 have increased much more sharply than EU output, while the development of the M1-ratio has been relatively stable. The figures also show that the movements in the liquidity ratios for the Divisia aggregates are – as found earlier by Fase (1985) for the Netherlands – fairly similar to those in the M1-ratio. This is plausible since M1-balances are primarily held for transaction purposes. Their developments are also fairly similar to those reflected by the modified M2 and M3-ratios. The relatively strong increase of the modified ratios at the end of the 1980s which, as argued before, is likely to be the result of the inverted yield curve, is not present in the Divisia monetary ratios. An explanation could be that Divisia aggregates, offering, like M1, a measure of the balances mainly held for transaction purposes, are not very sensitive to changes in the slope of the yield curve. A general conclusion following from Figures 4.1 and 4.2 is that according to the Divisia aggregates, monetary expansion has been fairly modest. This conclusion, which is in line with the results discussed before, indicates again that the strong increase of the M2 and M3-ratios is primarily due to portfolio investment considerations rather than to an expansionary monetary policy.

---

7 For an extensive discussion of the construction of the Divisia aggregates the reader is referred to e.g. Fase and Winder (1994).

Figure 4.1
Liquidity ratios for M1, M2, Divisia-M2 and modified M2 ratio

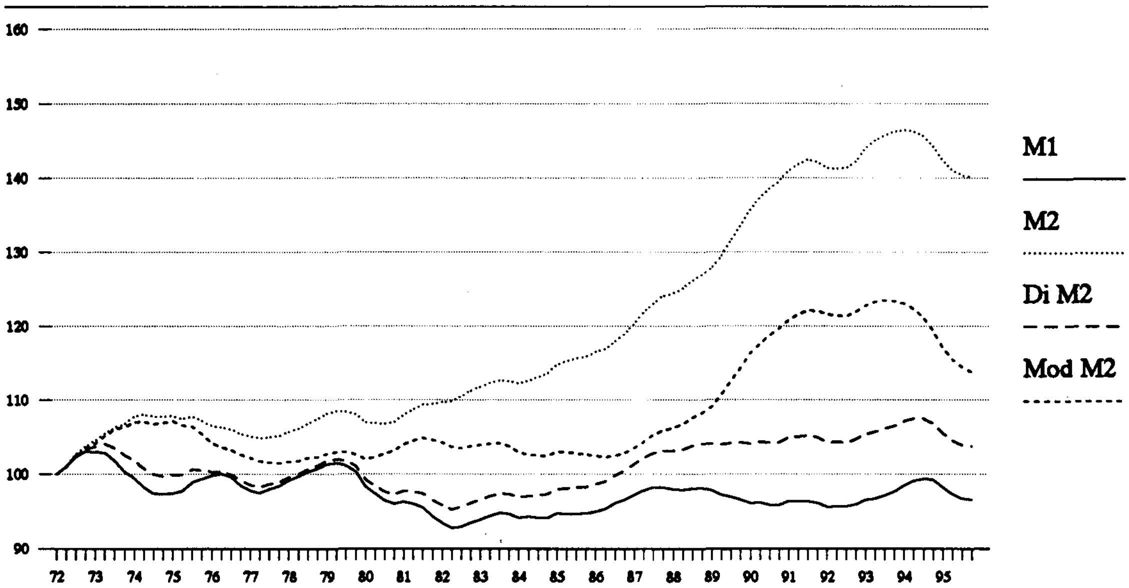
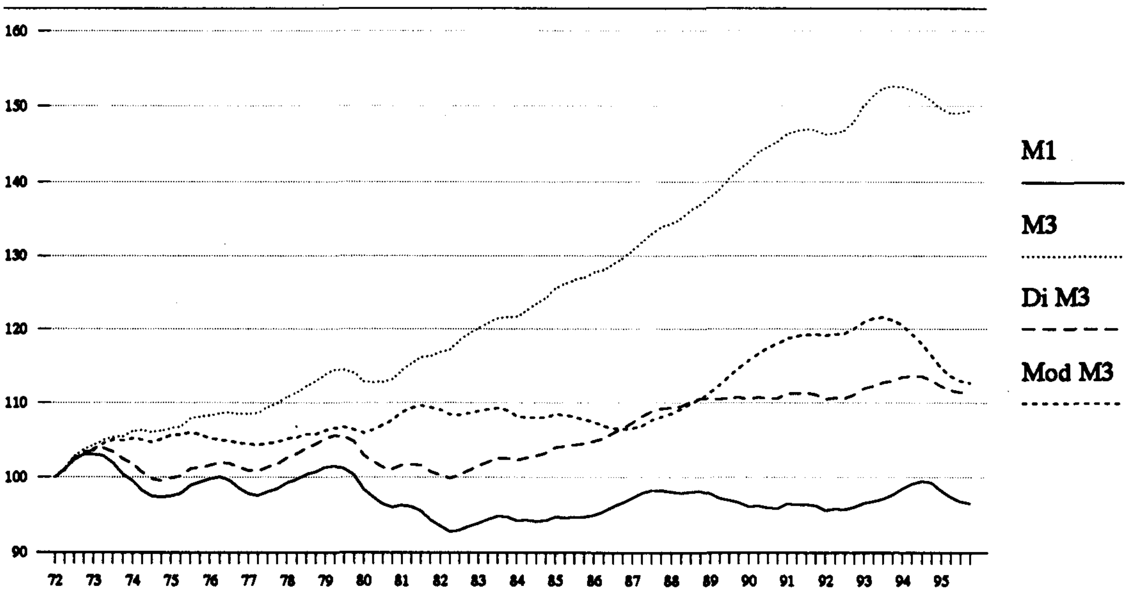


Figure 4.2
Liquidity ratios for M1, M3, Divisia-M3 and modified M3 ratio



Concluding remarks

This study presented the results of an empirical analysis of the demand for money in the European Union as a whole over the period 1971-95, with a particular focus on the impact of financial wealth. In order to obtain a comprehensive picture we did not restrict the analysis to just one monetary aggregate, but considered M1, M2 and M3. The empirical evidence shows a substantial impact of wealth on the demand for M2 and M3. In contrast, no influence of wealth on the demand for M1 could be found, indicating that portfolio investment considerations are less important for M1-balances. The empirical results show also that for M2 and M3 the income and wealth elasticities add to 1, and that M2 is more sensitive to changes in interest rates and inflation than M1 or M3. An extensive specification analysis indicated that the demand for the three money concepts considered is stable for the European Union as a whole. This is important because it is one of the main conditions for a policy of monetary targeting.

In the European Union, the broad monetary aggregates M2 and M3 have increased much more strongly than nominal output, in particular since the beginning of the 1980s. In the spirit of the monetarist view of inflation, according to which – loosely speaking – inflation is the result of too much money chasing too few goods, this monetary expansion should have led to an upward tendency in inflation. In fact, the 1980s have actually shown a substantial reduction of the inflation rates. The present analysis indicates that the strong growth of financial wealth in the 1980s offers an explanation of the sharp increase of the M2 and M3-ratios. Taking into account the development of wealth leads to the conclusion that the monetary expansion has been fairly modest, and thus that the strong increase of M2 and M3 should mainly be attributed to portfolio investment considerations.

Including wealth among the set of determinants of the demand for money is one way to incorporate portfolio investment considerations. An alternative approach is offered by the construction of Divisia aggregates. These money concepts correct for differences in the degrees of liquidity of the various liquid assets in order to provide a measure of the volume of liquidities held for transaction purposes. In this approach the growth of the broad monetary aggregates is thus corrected for the growth of the balances held for portfolio investment purposes. Our results for the Divisia aggregates for the European Union confirm also that monetary expansion has been modest, as the development of the liquidity ratios for Divisia aggregates are similar to that of the M1-ratio. This indicates that M1 is a reliable indicator of the volume of liquidities held for transaction purposes. In this regard, the informational content of the broader aggregates is frequently distorted as a result of portfolio investment considerations.

Though the development of financial wealth offers an explanation for the strong increase of the broad monetary aggregates, it cannot be ignored that the resources embodied in the wealth position of the private sector – as far as it concerns the household sector – is a source of future consumption. The developments since the beginning of the 1980s indicate that these cumulated resources have not been used for consumption on a large scale. For the near future, it is likely that this will not be the case either. Demographic factors, like the ageing of the population, and austerity policies of the governments, including a re-shaping of the social welfare system, may well be an impetus for further increases of wealth. The fact that wealth may be used for actual spending objectives is, however, a reason for the monetary authorities to monitor the development of the wealth position of the private sector. This applies also if wealth is not yet used for current consumption, but intended for future consumption. The results in our study show that wealth has a substantial impact on the demand for money. For an adequate assessment of the monetary developments it is, therefore, useful to take into account the development of wealth.

Appendix A: Further information on the data for financial wealth

This appendix contains further information on the data used in the analysis of demand for money in the European Union. First, we will discuss the construction of the data on wealth for the EU-countries and next, we will shortly give some information concerning the sensitivity of the EU-wide aggregates with respect to the various available aggregation methods.

Since data on financial wealth are not readily available for all countries we constructed a database consisting of balance sheets for four sectors, viz. the non-monetary private sector, money creating institutions (monetary authorities and deposit money banks, comprising commercial banks and other banks that accept demand deposits), the central government and the foreign sector. These balance sheets reflect the mutual claims and debt relationships between the various sectors and enable the calculation of net financial wealth of the private sector.

In order to achieve consistency between the definitions of the various debt instruments for the individual countries, we have used only data published in the International Financial Statistics (IFS) and, occasionally, in the Government Finance Statistics Yearbooks. The information reported by the national central banks and statistical agencies to the IMF concerns relatively uniform definitions of the various concepts. Though full congruence between the national concepts is not feasible due to specific national characteristics, the differences are likely to be small. In the following discussion we will focus on the balance sheets for the Netherlands, mentioning also the IFS codes of the variables used, which are reported on the country page for the Netherlands. For a number of countries some amendments to the Netherlands' balance sheets were necessary, but these are grosso modo of relatively little importance.

The balance sheet of the private sector reads as:

$$M1 + QM + CR_{ps,dmb} + CR_{ps,gov} + NFA_{ps} = CR_{dmb,ps} + W_{ps}$$

with $M1$ (IFS line 34) currency and demand deposits, QM (line 35) quasi-money, $CR_{ps,\,dmb}$ (line 36ab) and $CR_{ps,\,gov}$ claims of the private sector on deposit money banks and the government respectively, NFA_{ps} net foreign assets of private sector and $CR_{dmb,\,ps}$ (line 32d) claims of monetary authorities and deposit money banks on the private sector. Net financial wealth is the differential between total assets and $CR_{dmb,\,ps}$.

The balance sheet of the money creating institutions follows from the "monetary survey" on the country pages of the IFS, containing the consolidated monetary authorities' and deposit money banks' data:

$$CR_{ps,dmb} + CR_{cb+dmb,gov} + NFA_{cb+dmb} = M1 + QM + CR_{ps,dmb} + REST$$

with $CR_{cb+dmb,\,gov}$ (line 32an) net claims on central government, NFA_{cb+dmb} (line 31n) net foreign assets of monetary authorities and deposit money banks, $REST$ (line 37r - line 32b) a residual item comprising, among other things, claims on local government. Other debt instruments are defined above for the private sector.

The balance sheet for the central government reflects that total debt is financed abroad or domestically:

$$DEBT = CR_{dmb+cb,gov} + CR_{ps,gov} + CR_{f,gov}$$

with $DEBT$ (line 88a + line 89a) total debt of the central government and $CR_{f,\,gov}$ (line 89a) foreign government debt; the other debt instruments are already defined above.

The highly stylised balance sheet of the foreign sector reflects that the counterpart of net foreign assets of the private sector and the money creating institutions is the sum of foreign government debt and the cumulation of current account balances:

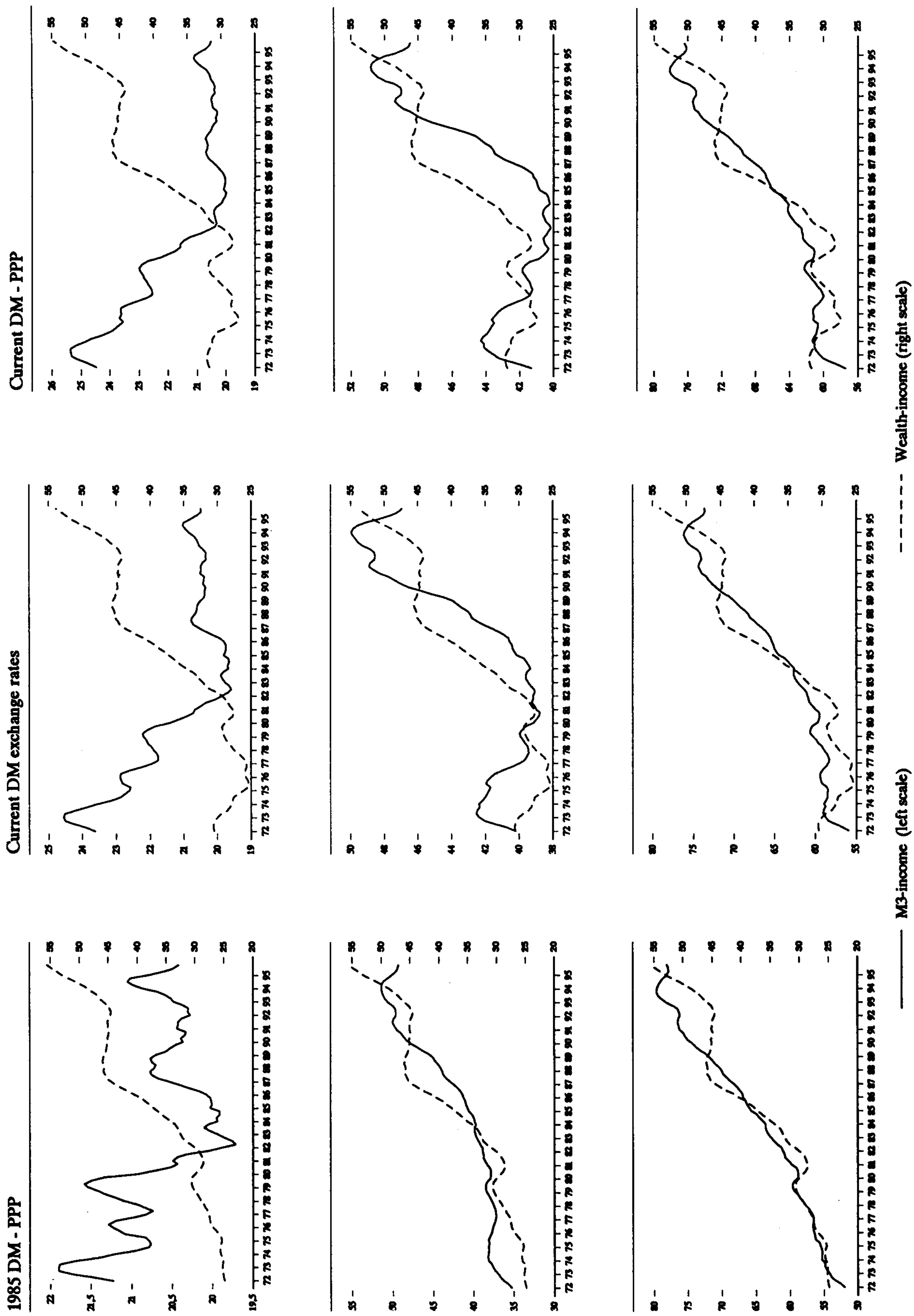
$$\sum CA + CR_{f,gov} = NFA_{cb,dmb} + NFA_{ps}$$

with CA (line 77a.d converted in the national currency with line rf) the current account balance and the other variables already defined above. In our calculations we cumulated current account balances since 1956, assuming a starting value of 0 for $\sum CA$ in 1955.

With the balance sheets for the money creating institutions, the central government and the foreign sector, the debt instruments of the balance sheet for the non-monetary private sector can be determined. $M1$, QM, $CR_{ps,\,dmb}$ and $CR_{dmb,\,ps}$ follow from the consolidated balance sheet for the monetary authorities and the deposit money banks. $CR_{ps,\,gov}$ can be derived from the balance sheet for the central government and NFA_{ps} follows from the stylised balance sheet for the foreign sector. Net financial wealth is next calculated as the differential between the total of assets of the private sector and $CR_{dmb,\,ps}$.

Section 1 of the main text discussed several options for constructing EU-wide data of the monetary aggregates, nominal output and wealth. In a tentative analysis of the influence of wealth on the demand for money, we depicted, in Figure 1, the development of the various liquidity ratios and the wealth-income ratio, showing that the rise of the liquidity ratios for M2 and M3 grosso modo coincides with the rise of the wealth-income ratio. This is prima facie evidence in favour of the hypothesis that the relatively strong increase of wealth is the driving force behind the monetary expansion as indicated by the growth of the liquidity ratios. In the main text we considered the EU-wide aggregates calculated with 1985 DM exchange rates as conversion factor. In order to assess whether these stylised facts are the result of using 1985 DM exchange rates to express the national variables in a common currency, we have also computed the EU-wide aggregates with the alternative conversion factors, viz. 1985 purchasing power parities against the DM and current DM exchange rate and DM purchasing power parities. The liquidity ratios and wealth-income ratio for the variables according to these alternative calculation methods are shown in Figure A1. The picture corresponds nicely with that in Figure 1, warranting the conclusion that the specific choice of the aggregation method, though influencing the results in their details, generally does not affect the overall pattern. As Figure A1 shows, the rather similar development of the liquidity ratios for M2 and M3 on the one hand and the wealth-income ratio on the other is no artefact of the specific conversion method.

Figure A1
Liquidity ratios and wealth-income ratios



Appendix B: Estimation results

This appendix presents the estimated demand for money equations for the European Union as a whole. The notation is the same as in the main text. Since the data are not seasonally adjusted, seasonal dummies s_1, s_2 and s_3 are included in the estimation equations. Moreover, we included a dummy variable D_{903} with the value 1 in 1990Q3 and 0 otherwise to account for the effect of the German unification. The t-values of the estimates, reported between brackets in the following equations, are adjusted for the possible presence of heteroscedasticity. $\overline{R}^2$ denotes the multiple correlation coefficient adjusted for degrees of freedom and SE is the standard error of the residuals.

$$\Delta \ln(M1/P) = -0.1833\left(\ln(M1/P)_{-1} - \ln y\right) - 0.0253 \ln y - 0.0016 rs_{-1} - 0.0010 rl_{-1} - 0.0010 \dot{p}$$

(3.19) (2.90) (2.16) (2.17) (2.17)

$$-0.0064 \Delta conj_{-2} - 0.0032 \Delta rs_{-1} - 0.0049 \Delta \dot{p} + 0.2648 \Delta \ln(M1/P)_{-4} + 0.1457 \Delta \ln y_{-4}$$

(2.57) (2.46) (3.49) (2.96) (2.74)

$$+0.1850 - 0.0004 s_1 + 0.0072 s_2 + 0.0048 s_3 + 0.0202 D_{903}$$

(2.69) (0.04) (1.51) (0.59) (9.80)

Estimation period: 1972Q1-1995Q4. $\overline{R}^2 = 0.89$. SE $= 0.0067$.

$$\Delta \ln(M2/P) = -0.0866\left(\ln(M2/P)_{-2} - \ln y_{-1}\right) + 0.0218\left(\ln w - \ln y_{-1}\right) + 0.0011 rs - 0.0017 rl_{-2}$$

(5.28) (2.58) (1.87) (3.33)

$$-0.0017 \dot{p}_{-2} - 0.0052 \Delta conj_{-3} - 0.0058 \Delta rl_{-1} + 0.2355 \Delta \ln(M2/P)_{-4} + 0.2430 \Delta \ln y$$

(3.33) (2.95) (2.92) (3.07) (3.70)

$$-0.0982 \Delta \ln w_{-3} + 0.0533 + 0.0114 s_1 + 0.0042 s_2 + 0.0117 s_3 + 0.0236 D_{903}$$

(2.10) (4.53) (1.26) (0.72) (1.40) (9.84)

Estimation period: 1972Q1-1995Q4. $\overline{R}^2 = 0.68$. SE $= 0.0080$.

$$\Delta \ln(M3/P) = -0.1032\left(\ln(M3/P)_{-2} - \ln y\right) + 0.0352\left(\ln w_{-1} - \ln y\right) + 0.0011 rs - 0.0014 rl$$

(5.77) (3.55) (2.14) (3.09)

$$-0.0014 \dot{p} - 0.0040 conj_{-1/2} - 0.0034 \Delta(rl - rs)_{-2} + 0.1805 \Delta \ln y_{-4} + 0.1121 \Delta \ln w$$

(3.09) (4.70) (2.58) (3.26) (3.99)

$$-0.0634 \Delta \ln w_{-3} + 0.0943 + 0.0228 s_1 + 0.0034 s_2 + 0.0121 s_3 + 0.0331 D_{903}$$

(1.92) (5.94) (2.95) (0.72) (1.65) (16.16)

Estimation period: 1972Q1-1995Q4. $\overline{R}^2 = 0.77$. SE $= 0.0058$.

Appendix C: Misspecification analysis

This appendix assesses the relationships presented in Appendix B. A first issue concerns the restrictions imposed on the estimated parameters, of which two categories can be distinguished. The first one concerns restrictions on the parameters of the long-run demand for money relationship. Examples are the restriction that the price elasticity equals unity or the restriction that the sum of the income and wealth elasticities is equal to unity. In the estimated equations we have also imposed the restriction that the long-run semi-elasticities with respect to the long-term interest rate and the inflation rate are equal. All these restrictions constitute testable hypotheses, for which the relevant t-statistics are reported in Table C1. In all cases, the restrictions imposed are justified from a statistical point of view.

The second class of restrictions concerns the parameters describing the short-run adjustment processes. An example is the error correction term in the equation for M1, $\left(\ln(M/P)_{-1} - \ln y\right)$, which arises from imposing a restriction on the coefficients of the familiar error correction term $\left(\ln(M/P)_{-1} - \ln y_{-1}\right)$, $\ln y_{-1}$ and $\Delta \ln y$. An alternative example is the restriction on the coefficients of Δrs_{-2} and Δrl_{-2} in the equation for M3, implying that the change of the slope of the yield curve rather than separate changes of Δrs and Δrl impacts on the demand for M3. These restrictions on the parameters describing the short-run dynamics have been tested as well. The results are not reported, but again all the restrictions imposed are valid from a statistical point of view.

In our statistical evaluation we considered next the properties of the residuals. We tested for residual autocorrelation, normality of the disturbance terms and heteroscedasticity of the ARCH-type. The null hypothesis of white noise errors has been tested with the Ljung-Box test statistic and the F-version of the Lagrange Multiplier test statistic.

Table C1

Statistics for restrictions on long-run parameters

Monetary aggregate	**Elasticities**			**Semi-elasticities**	
	$\alpha_1 = 1$	$\alpha_2 + \alpha_3 = 1$	$\alpha_3 = 0$	Restriction imposed	t-statistic
M1[1]	0.53	2.90*	0.03	$\alpha_5 = \alpha_6$	0.83
M2	1.25	0.33	2.58*	$\alpha_5 = \alpha_6$	1.82
M3	0.17	1.52	3.55*	$\alpha_5 = \alpha_6$	1.42

Note: Long-run money demand $M = P^{\alpha_1} y^{\alpha_2} w^{\alpha_3} \exp(\alpha_4 rs + \alpha_5 rl + \alpha_6 \dot{p})$; values of t-statistics are given;

* indicates significance at a significance level of 0.05.

[1] $\alpha_2 + \alpha_3 = 1$ corresponds to the restriction that the income elasticity equals 1.

As an alternative hypothesis we considered serial autocorrelation of order p, with p=1, 4, 8 and 12. The null hypothesis of normality of the residuals has been examined with the Jarque-Bera statistic and homoscedasticity has been tested against the alternative hypothesis of ARCH(p) processes, p=1 and 4, with the F-version of the familiar Lagrange Multiplier statistic. The results are reported in table C2 together with the relevant distribution under the null hypothesis and show, with the exception of the Ljung-Box statistic for 12th order residual autocorrelation in the case of M3 and the Jarque-Bera statistic in the case of M2, that all statistics are insignificant at a significance level of 0.05. In Section 3 of the main text we considered as an alternative test of misspecification the Chow-statistics for the null hypothesis of parameter stability for the periods before and after the German unification. We examined also the CUSUM and CUSUMSQ statistics based on the recursive residuals. This stability analysis yielded satisfactory results for the specified equations, and, together with the results reported in Table C2, warrants the conclusion that the specified equations are appropriate from a statistical point of view.

Table C2
Results of misspecification analysis

Source of misspecification/ test-statistic		**Distribution**	**M1**	**M2**	**M3**
Residual autocorrelation					
Ljung-Box Q-statistic	1	$\chi^2(1)$	0.22	2.46	0.33
	4	$\chi^2(4)$	3.90	4.07	8.65
	8	$\chi^2(8)$	5.23	12.87	12.73
	12	$\chi^2(12)$	7.44	15.21	21.06*
LM F-statistic	1	F(1,81)	0.23	2.11	0.29
	4	F(4,78)	1.11	1.21	2.22
	8	F(8,74)	0.76	1.51	1.38
	12	F(12,70)	0.60	1.14	1.77
Normality					
Jarque-Bera statistic		$\chi^2(2)$	0.18	21.61*	1.17
Heteroscedasticity-ARCH					
LM F-statistic	1	F(1,93)	0.00	0.44	1.94
	4	F(4,87)	1.17	0.45	1.23

Note: * significant at a significance level of 0.05.

The long-run relationships implied by the dynamic demand for money equations, have been analysed by means of cointegration tests. One way to examine cointegration is to apply Johansen-type tests. Since these tests require the specification of a VAR-model rather than a single equation approach, the Johansen procedure has not been applied. Instead, we have investigated whether the deviations of the long-run relationship follow a stationary process. Since the data used are not seasonally adjusted, the procedure put forward by Hylleberg, Engle, Granger and Yoo (1990) has been applied. Letting x_t denote the deviations from the long-run relationship, $y_{1t} = (1+L+L^2+L^3)x_t$, $y_{2t} = -(1-L)(1+L^2)x_t$, $y_{3t} = -(1-L)(1+L)x_t$, L the lag-operator, and ε_t the disturbance term, the equation has been estimated such that:

$$\Delta_4 x_t = \pi_1 y_{1t-1} + \pi_2 y_{2t-1} + \pi_3 y_{3t-2} + \pi_4 y_{3t-1} + \varepsilon_t \qquad \text{(C.1)}$$

The restriction $\pi_1 = 0$ corresponds to a unit root equal to 1, while the restriction $\pi_2 = 0$ implies that a unit root $L = -1$ is present; π_3 and π_4 are both equal to zero, the unit roots $L = \pm i$ are present. As these unit roots are each other's complex conjugates both π_3 and π_4 must be equal to zero. These restrictions can be tested by means of the t-values of π_i, i=1,..,4 and the F-test for the hypothesis $\pi_3 = \pi_4 = 0$.

Since the conventional statistical theory does not apply to these unit root tests, the critical values derived by Hylleberg et al. have been used.

Table C3

Results of cointegration analysis

Monetary aggregate	$t(\pi_1)$	$t(\pi_2)$	$t(\pi_3)$	$t(\pi_4)$	$F(\pi_3,\pi_4)$
M1	-2.89*	-4.47**	-6.95**	-2.85**	37.31**
M2	-2.90*	-5.27**	-6.76**	-6.74**	87.01**
M3	-2.69*	-4.79**	-6.13**	-7.12**	80.01**

Note: * (**) significant at a significance level of 0.10 (0.01).

For the cointegration test, equation (C.1) was estimated with seasonal dummies, a constant term and the first four lags of $\Delta_4 x_t$ included as additional explanatory variables to eliminate possible residual auto-correlation. Subsequently, the most insignificant lag of $\Delta_4 x_t$ was removed and repeating this procedure ultimately yields an equation with only significant lags of $\Delta_4 x_t$. The results for the t and F-test statistics for this ultimate equation are shown in Table C3. They show that both for M1, M2 and M3 the hypothesis that a unit root $L=1$ is present cannot be rejected at a 5% significance level. Using a 10% significance level, however, cointegration at the long-run frequency can be accepted. The results further indicate that seasonal unit roots are not present at the 5% significance level. An alternative test for cointegration has been put forward by Kremers et al. (1992). This test is based on the t-statistic of the estimated coefficient of the error correction term in the dynamic equation. Under specific assumptions this statistic is normally distributed. If these assumptions are violated, the Dickey-Fuller critical values can be used. Using this procedure, the results reported in Appendix B yield supportive evidence for the cointegration hypothesis.

References

Barnett, W. A. (1978): "The user cost of money". *Economics Letters*, No. 1, pp. 145-149.

Bekx, P. and G. Tullio (1989): "A note on the European Monetary System and the determination of the DM-dollar exchange rate". *Cahiers économiques de Bruxelles*, No. 123, pp. 329-343.

Brainard, W. C. and J. Tobin (1968): "Pitfalls in financial model building". *American Economic Review*, No. 58, pp. 99-122.

Butter, F. A. G. and M. M. G. Fase (1981): "The demand for money in EEC countries". *Journal of Monetary Economics*, No. 7, pp. 201-230.

Deutsche Bundesbank (1995): "Uberprüfung des Geldmengenziels und Neuordnung der Mindestreserve". *Monatsberichte*, July.

Divisia, F. (1925, 1926): "L'indice monétaire et la théorie de la monnaie". *Revue d'Economie Politique*, No. 39, pp. 842-861, pp. 980-1008, pp. 1121-1151; and No. 40, pp. 49-87.

Fase, M. M. G. (1985): "Monetary control: the Dutch experience – some reflections on the liquidity ratio", in C. van Ewijk and J. J. Klant (eds.), *Monetary conditions for economic recovery*, Dordrecht/Boston, pp. 95-125.

Fase, M. M. G. (1994): "In search for stability: an empirical appraisal of the demand for money in G7 and EC countries". *De Economist*, No. 142, pp. 421-454.

Fase, M. M. G. (1995): "Divisia aggregates and the demand for money in core EMU". *Onderzoeksrapport WO&E*, No. 424.

Fase, M. M. G. and J. B. Kuné (1974): "De vraag naar liquiditeiten in Nederland – 1952-1971". *De Economist*, No. 122, pp. 326-356.

Fase, M. M. G. and C. C. A. Winder (1993): "The demand for money in The Netherlands and other EC countries". *De Economist*, No. 141, pp. 471-496.

Fase, M. M. G. and C. C. A. Winder (1994): "Money demand within EMU – an analysis with the Divisia measure". *De Nederlandsche Bank Quarterly Bulletin*, pp. 25-55, September.

Fase, M. M. G. and C. C. A. Winder (1996): "Wealth and the demand for money: empirical evidence for The Netherlands and Belgium". *De Economist*, No. 144, forthcoming.

Filosa, R. (1995): "Money demand stability and currency substitution in six European countries (1980-1992)". *BIS Working Paper*, No. 30.

Friedman, M. (1971): "A theoretical framework for monetary analysis". *National Bureau of Economic Research*, Colombia University Press, New York.

Hall, S. G., S. G. B. Henry and J. B. Wilcox (1989): "The long-run determination of the UK monetary aggregates". *Bank of England Discussion Papers*, No. 41.

Hylleberg, S., R. F. Engle, C. W. J. Granger and B. S. Yoo (1990): "Seasonal integration and cointegration". *Journal of Econometrics*, No. 44, pp. 215-238.

Kremers, J. J. M., N. R. Ericsson and J. J. Dolado (1992): "The power of cointegration tests". *Oxford Bulletin of Economics and Statistics*, No. 54, pp. 325-348.

Kremers, J. J. M. and T. D. Lane (1990): "Economic and monetary integration and the aggregate demand for money in the EMS". *IMF Staff Papers*, No. 37, pp. 777-805.

Monticelli, C. and M. Strauss-Kahn (1992): "European integration and the demand for broad money". *BIS Working Paper*, No. 18.

Sterken, E. (1992): "The financial transactions motive in the Dutch money demand equation". *Kredit und Kapital*, No. 25, pp. 386-405.

Comments on: "Wealth and the demand for money in the European Union" by Martin M. G. Fase and Carlo C. A. Winder

by Benjamin H. Cohen

It has been reported that monetary targeting is among the policy procedures being considered for the new European central bank. If this is the case, then stable, reliable models of Europe-wide money demand will certainly be needed. In view of the breakdowns in the relationships between money and nominal variables that have been observed in the US and elsewhere, one might have grounds for scepticism regarding whether such models can be found for Europe.

This paper addresses these concerns by presenting plausible error-correction models for quarterly growth in the real money stock in the area of the current (15-member) EU over 1972-95. The authors obtain adjusted R-squared's of 0.89, 0.68, and 0.77 for M1, M2 and M3 respectively. Though most of the restrictions make sense, it would perhaps be helpful to see what value the various elasticities take when the restrictions assumed by the authors do not take hold. Some of the restrictions could be justified better, particularly setting the sum of the wealth and income elasticities to one and forcing the coefficients on inflation and the long-term interest rate to be equal. However, it is reassuring that the models pass a range of stability and specification tests and that they do not seem very sensitive to the method used for aggregating across currencies.

The key element in this good performance, at least for the M2 and M3 models, is probably the inclusion of the ratio of financial wealth to income among the independent variables. This is justified as a way to incorporate portfolio motives for holding short-term money-market instruments. The authors support this by showing that Divisia indices for M2 and M3, which weight monetary components by liquidity, behave more or less similarly to M1 over the period studied. The authors attribute price stability in the 1980s, when broader monetary aggregates were rising, to the stability of wealth-adjusted money demand, though it must be said that in their graphs (Figure 3) wealth-adjusted money looks pretty stable in the 1970s as well.

In their conclusion, the authors suggest that this relationship is grounds for European monetary authorities to pay close attention to wealth accumulation. The question of whether excessive asset-price run-ups can be inflationary is certainly an interesting and timely issue. I would contend, however, that if anything the present paper offers evidence that wealth accumulation is *not* inflationary: the authors show that substantial increases in certain categories of liquid assets merely reflect portfolio decisions in the context of rising wealth, and not a build-up of the kind of liquidity pressures that are thought to be inflationary.

Inflation and economic growth: some evidence for the OECD countries[1]

Javier Andrés and Ignacio Hernando

Introduction

During the last decade, inflation control has become the main goal of monetary policies in western economies. This move in monetary policy-making is firmly rooted in the belief, shared by many economists as well as politicians, that the costs of inflation are non-negligible, whereby keeping inflation under control pays off in terms of higher per capita income in the future.

The lack of theoretical models explicitly addressing the issue of the long-run effects of inflation has not prevented many researchers from trying to estimate inflation costs. The evidence so far is not conclusive. A series of recent papers have tried to assess the long-run impact of current inflation within the framework of the so called *convergence equations*. These equations can be derived from a theoretical model of economic growth, although the reasons for including the inflation rate among the determinants of growth remain somewhat unclear. This paper adheres to this approach to estimating the impact of inflation upon the long-run performance of the OECD countries. The approach has several advantages as compared with more standard ones. First, and foremost, an explicit model prevents the omission of relevant variables. Second, convergence equations allow for a variety of effects of inflation including those which reduce accumulation rates and those which undermine the efficiency with which productive factors operate. Finally, in this framework a clear distinction can be made between *level* and *rate of growth* effects of inflation; this difference matters as regards the size and the timing of the costs of inflation. The methodology also has some shortcomings. First, growth models focus on long-run issues, disregarding the short-run costs associated with disinflation (*the sacrifice ratio*). Second, the use of multi-country data sets imposes too many restrictions on the parameters to prevent a shortage of degrees of freedom. Also, the direction of causality among the variables included in convergence equations is not unambiguous. This paper tries, in several ways, to overcome these limitations to check the robustness of the inflation-growth empirical link.

The rest of the paper is organised as follows. Section 1 briefly summarises the literature dealing with the cost of inflation; the empirical model and the data used are also discussed in some detail. In Section 2 we present the convergence equations augmented with the rate of inflation. In Section 3, cross-country heterogeneity is allowed for in the convergence model, whereas in Section 4 standard causality tests are applied to the inflation-growth relationship. The final section concludes with some additional remarks. The main results of the paper can be summarised as follows. Even low or moderate inflation rates (as the ones we have witnessed within the OECD) have a negative but temporary impact upon long-term growth; this effect is significant and generates a permanent reduction in the level of per capita income. Inflation not only reduces the level of investment but also the efficiency with which productive factors are used. The estimated cost of a 1% rise in the inflation rate is a reduction, during rather long periods, of the annual growth rate of about 0.06%; in the long-run this leads to a reduction in the steady-state per capita income of about 2%. This result holds across different sub-samples (even excluding high-inflation countries) and is also robust to alternative econometric specifications. In particular, inflation Granger-causes income and the current and lagged

---

[1] This paper was presented at the NBER Conference on "The Costs and Benefits of Achieving Price Stability", Federal Reserve Bank of New York (February 20-21, 1997). A substantially revised version of the paper will be published at the Conference volume: "The Costs and Benefits of Achieving Price Stability" (edited by Martin Feldstein). We are grateful to Palle Andersen, Sean Craig, Juanjo Dolado, Rafael Doménech, Angel Estrada, Frederic Mishkin, Teresa Sastre, Javier Vallés and José Viñals for their comments and to Francisco de Castro for his excellent research assistance. The authors are member of the Banco de España, and Javier Andrés is, in addition, affiliated to the University of Valencia.

correlation between these two variables remains significant when we control for country-specific variables (such as the accumulation rates) and time invariant effects.

1. The theoretical framework

1.1 International evidence

The negative effects of inflation have been studied in the context of the models of economic growth (Orphanides and Solow (1990), De Gregorio (1993) and Roubini and Sala-i-Martín (1995)). The continuous increase of per capita income is the outcome of capital accumulation and the continuous improvement in the efficiency with which productive factors are used. The uncertainty associated with a high and volatile unanticipated inflation has been found to be one of the main determinants of the rate of return of capital and investment (Bruno (1993) and Pindyck and Solimano (1993)). But even fully anticipated inflation may reduce the rate of return of capital given the non-neutralities built into most industrialised countries' tax systems (Jones and Manuelli (1993) and Feldstein (1996)). Besides, high and volatile inflation undermines the confidence of foreign investors about the future course of monetary policy. Inflation also affects the accumulation of other determinants of growth such as human capital or investment in R+D; this channel of influence constitutes what is known as the *accumulation or investment effect* of inflation on growth.

But, over and above these effects, inflation also worsens the long-run macroeconomic performance of market economies by reducing the efficiency with which factors are used. This latter channel, also known as the *efficiency channel*, is harder to formalise in a theoretical model;[2] nonetheless, it is widely agreed that its importance in the transmission mechanism from inflation towards lower growth cannot be denied. A high level of inflation induces frequent changes in price lists which may be costly for firms (*menu costs*) and reduces the optimal level of cash holdings by consumers (*shoe-leather costs*). It also induces bigger forecast errors by distorting the information content of prices, encouraging economic agents to spend more time and resources on gathering information and protecting themselves against the damages caused by price instability, hence endangering the efficient allocation of resources.

Many authors have found a negative correlation between growth and inflation. The following paragraphs sum up the most significant features of several of these studies. Kormendi and Meguire (1985) estimate a growth equation with cross-section data and find that the effect of inflation on the growth rate is negative, although it loses explanatory power when the rate of investment is also included in the regression. This would indicate that the effect of inflation mainly manifests itself in a reduction in investment but not in the productivity of capital. Grier and Tullock (1989) estimate a model that excludes the rate of investment and includes several measures of nominal instability (inflation rate, price acceleration and standard deviation of inflation). The results differ according to the group of countries in question, but for the OECD only the variability of inflation seems to have a significant and negative effect on growth.

Starting from these seminal works, the study of the long-run influence of inflation has primarily developed within the framework of convergence equations drawn from economic growth theory.[3] Fischer (1991, 1993) detects a significant influence of several short-term macroeconomic

---

2 As Briault (1995) has rightly pointed out, it is very difficult to derive a significant effect of inflation on factor productivity in frictionless general equilibrium competitive models.

3 The next section develops these equations and discusses their properties. Several exceptions, however, are worth noting: the studies of Grimes (1991) for the OECD, Smyth (1994) for the United States, Cardoso and Fishlow (1989)

indicators on the growth rate. Inflation reduces both capital accumulation and total factor productivity. Cozier and Selody (1992) find that, for the sub-sample of OECD countries, inflation affects the level rather than the growth rate of productivity, whereas the variability of inflation does not seem to have any appreciable effect. This finding coincides with the result obtained more recently by Barro (1995) for a sample of 120 countries, who reports a negative long-run effect of inflation,[4] which is more pronounced at higher levels of the inflation rate. The general conclusion of these and other studies (De Gregorio (1992a, 1992b and 1994) and Motley (1994)) is consistent with the negative correlation between inflation and income in the long run suggested in the theoretical literature. However, the consensus in this respect is far from absolute, and several authors have criticised these findings, arguing that the lack of a fully developed theoretical framework makes it difficult to interpret the empirical correlations and that these are not robust to changes in the econometric specification. The latter argument is developed in Levine and Renelt (1992), Levine and Zervos (1993) and Clark (1993). Levine and Renelt carry out an exhaustive sensitivity analysis of the set of explanatory variables contained in the literature on economic growth, showing how the effect of most of these variables (inflation among them) is not invariant to changes in the information set on which this effect depends. Nor do these results, in turn, escape criticism. Sala-i-Martín (1994) argues that the problem of finding a macroeconomic variable, the effect of which is invariant to alternative specifications of the convergence equation, should not be taken to mean that this influence is absent, but should instead be viewed as a sign of the difficulty of finding indicators that can adequately capture this effect for any period and group of countries. Lastly, Andrés, Doménech and Molinas (1996b) show that, for the OECD as a whole, the variables of macroeconomic policy are even more robust than the rates of accumulation in explaining economic growth.

1.2. The effects of inflation in a neoclassical growth model

There are a number of advantages to estimating the correlation between inflation and growth within the framework of the convergence equations proposed by Barro and Sala-i-Martín (1991), as these represent the main empirical approach to growth models with constant returns.[5] Let us consider a growth model (Mankiw, Romer and Weil (1992)) in which technology is represented by the following production function with constant returns ($\alpha + \beta + \gamma = 1$),

$$Y_t = \left(A_t L_t\right)^{\beta} K_t^{\alpha} H_t^{\gamma} \quad (1)$$

Total factor productivity (A_t) grows at the constant exogenous rate ϕ, whereas fixed capital (K) and human capital (H) grow in proportion to the output assigned for their accumulation.[6] Let us also assume that the depreciation rates of both factors are the same. With these assumptions, it is possible to arrive at the following equation of growth in per capita income between two moments in time (t, $t + \tau$):

$$y_{T+\tau} - y_T = \phi\tau + \left(1 - e^{-\lambda\tau}\right)\left[\Omega^C + y_T^* - y_T\right] \quad (2)$$

who use a panel of five-year averages for 18 Latin American countries, Burdekin, Goodwin, Salamun and Willett (1994) and Bruno (1993). In all these studies, a significant negative effect of inflation on growth is reported.

4 Whereas the effect of the variability of inflation is not invariant to alternative specifications.

5 De Gregorio (1993) and Roubini and Sala-i-Martín (1995) provide more elaborate models of the interaction between inflation and growth.

6 In the original formulation of Solow (1956), the rate of technological progress was exogenous, while in more recent models it can be explained by the set of resources assigned to research, market size, learning-by-doing, etc.

where y represents the logarithm of per capita income in the periods indicated by the subscripts, and y^* represents its stationary state value. According to equation (2), the growth rate of an economy will have a component determined by the growth in factor productivity at a rate ϕ and another resulting from the economy's propensity to move towards its steady-state level if, for some reason (shocks, initial conditions, etc.), it lies outside. λ is the rate at which the economy closes the gap between its current income level and its potential or steady-state level.[7] The latter is, in turn, determined by the parameters of the production function and by the rates of accumulation of the productive factors in the stationary state:

$$y_T^* = \Omega^s + \phi T + \beta^{-1}\left[\alpha s_{Tk}^* + \lambda s_{Th}^* - (\alpha+\gamma)\log\left(n_t^* + \phi + \delta\right)\right] \tag{3}$$

where s_k^* is the logarithm of the rate of investment, s_h^* represents the logarithm of the rate of accumulation of human capital, and n^* is the growth rate of the population, all evaluated at their steady-state level; lastly, δ is the depreciation rate of capital (physical and human). This we will assume to be exogenous and equal to an annual 3%,[8] while Ω^c and Ω^s are two constants that combine different parameters of the model and the starting level of technology (A_T).

The use of equations (2) and (3) as the analytical framework does not presuppose the acceptance of the exogenous growth model as the only possible representation of the behaviour of OECD economies in the long run. The main advantage of this model is that it systematically captures most of the factors that the literature on economic growth has pointed to as determinants of growth; this reduces the risk of omitting relevant regressors entailed in ad hoc specifications.[9] To test the influence of inflation on income in the long run the usual procedure (Cozier and Selody (1992)) is to augment equations (2) and (3), by assuming that the productivity index (A_T) evolves in accordance with expression (4), which reflects the influence of the inflation rate (π) and its variability (σ):

$$A_t = A_0 \exp(\phi t)\exp(\mu_1 \pi)\exp(\mu_2 \sigma) \tag{4}$$

The system of equations to be estimated is thus one formed by equations (2) and (3'):

$$y_T^* = \Omega^s + \phi T + \mu_1 \pi_T + \mu_2 \sigma_T + \beta^{-1}\left[\alpha s_{Tk}^* + \gamma s_{Th}^* - \left(\overset{\circ}{\alpha+\gamma}\right)\log\left(n_T^* + \phi + \delta\right)\right] \tag{3'}$$

This simple structure allows us to test the different hypotheses considered in this paper. First, the presence of the rates of factor accumulation in (2) and (3') is useful to discriminate between the two channels through which macroeconomic distress can influence the growth rate. Thus, if inflation influenced growth solely through its direct impact on total factor productivity, we could expect the coefficient μ_1 estimated in equations (2) and (3') to be independent of the rates of accumulation. In contrast, this coefficient varies substantially, we can conclude that there is an inflation effect on agents' investment efforts.[10] Second, the exogenous growth model specifies the

---

7 This rate can be written as: $\lambda = (1-\alpha-\gamma)\left(n^* + \phi + \delta\right)$.

8 To use a value that is standard in the literature.

9 Unlike growth equations that do not include the catching-up component, the convergence equation provides a way of controlling the level of per capita income when analysing the determinants of its growth rate.

10 In this case, the possible impact of inflation on long-run growth should be evaluated by estimating the investment equations.

determinants of the level of per capita income in the long run and also of the sustained growth rate. Inflation can affect the one and/or the other, although the implications in terms of welfare are different. According to equation (4), the impact of inflation basically manifests itself in the potential level of per capita income, but not in sustained growth (represented by ϕ). To examine the latter possibility, we will later consider an alternative specification, (4'), which allows inflation to influence the long-run growth rate:[11]

$$A_t = A_0 \exp[(\phi + \phi'\pi)t]\exp[\mu_1\pi]\exp[\mu_2\sigma] \tag{4'}$$

resulting in a new system of equations being estimated as:

$$y_{T+\tau} - y_T = (\phi + \phi'\pi)\tau + \left(1 - \bar{e}^{\lambda\tau}\right)\left[\Omega^c + y_T^* - y_T\right] \tag{5}$$

$$y_T^* = \Omega^s + (\phi + \phi'\pi)T + \mu_1\pi + \mu_2\sigma_T + \beta^{-1}\left[\alpha s_{Tk}^* + \gamma s_{Th}^* - \left(\overset{\circ}{\alpha + \gamma}\right)\log\left(n_T^* + \phi + \delta\right)\right] \tag{6}$$

The specification can be also used to test for possible non-linear effects of inflation on income. A linear version of the convergence equation is given by:

$$y_{T+\tau} - y_T = \psi_0 + \psi_1 T + \psi_2 y_T + \psi_3 s_{TK}^* + \psi_4 s_{TH}^* + \psi_5 \log\left(n_T^* + \phi + \delta\right) + \psi_6 \pi_T + \psi_7 \sigma_T \tag{5'}$$

This approach also has some drawbacks. Cross-country convergence regressions are often blamed for two misspecification flaws: the ambiguous pattern of causality among the variables included in the equations and the absence of a proper account of country specifics. We take account of these shortcomings in Sections 3 and 4. In Section 2 we estimate the elasticity of growth with respect to inflation in a standard convergence equation. We use four-year averages of OECD annual data; for a detailed discussion of the features of the data base, see Dabán, Doménech and Molinas (1996).

2. Estimation of the effect of inflation

Tables 1 and 2 present the results of estimating the steady-state and convergence equations derived from the exogenous growth model. In order to reduce the risk of simultaneity biases, all equations reported have been estimated by instrumental variables, using one and two-periods-lagged regressors as instruments. The results seem to be quite robust, both in the linear and in the non-linear specifications. Linear models (equation (5')) are presented in Table 1. In columns 1 and 2 we estimate different versions of the convergence equation whereas in columns 3 and 4 we present estimates of the steady-state equation. The evidence in columns 1 and 2 is consistent with the convergence property implied by the Solow growth model. The parameter for initial per capita income is negative and highly significant both when steady-state variables are included (conditional convergence) and when they are not (unconditional convergence). In column 2, the estimated coefficients for the input accumulation rates have the expected sign, although the one for human capital is non-significant. The estimated parameter for the trend, which according to the theoretical model is proxying technological progress, has an unexpected negative sign.[12] On the other hand, the

---

[11] This is the specification proposed by Motley (1994), from which the effect of the variability of inflation is excluded in order to simplify the expression.

[12] One reason for this result is that the trend may be capturing the process of sustained reduction in the rate of growth of per capita income suffered by OECD countries during part of the sample period. As an alternative approach, time

estimated coefficient for the trend has the expected positive sign in the steady-state equation (columns 3 and 4). However, the values of the estimated parameters for the input accumulation rates merit caution in interpreting the results.

Table 1

Linear models (equation (5'))

Dependent variable:	(1) Δy	(2) Δy	(3) y^*	(4) y^*
ψ	0.350	-0.275	2.304	-1.527
	(8.98)	(1.21)	(25.50)	(1.39)
ψ_1	-0.010	-0.010	0.077	0.044
	(3.33)	(3.02)	(5.27)	(2.97)
ψ_2	-0.074	-0.090	-	-
	(4.60)	(5.12)		
ψ_3	-	0.044	-	0.169
		(1.95)		(1.54)
ψ_4	-	0.026	-	0.771
		(0.82)		(5.64)
ψ_5	-	-0.143	-	-0.000
		(1.91)		(0.00)
ψ_6	-0.0019	-0.0012	-0.023	-0.012
	(2.98)	(1.71)	(8.25)	(3.75)
R^2	0.31	0.39	0.38	0.55
σ	0.055	0.053	0.298	0.258

Notes: Estimation method: instrumental variable. Instruments: constant, trend and first and second-order lags of the dependent variable and of the regressors.

When the inflation rate and its variability (proxied by the Pearson coefficient) are included, both individually and jointly, the rest of the parameters do not change significantly. When both variables are jointly included, the coefficient for the inflation rate is negative and significant, both in the convergence and in the steady-state equation, whereas no significant effect is found for the variability of inflation. Thus, the equations presented hereafter only include the inflation rate. When the factor accumulation rates are included (columns 2 and 4) the size of the inflation effect is smaller than when they are omitted (columns 1 and 3), but it is still significant. This suggests that there are two channels by which inflation influences growth: first, through a reduction in the propensity to invest, and second, through a reduction in the efficiency in the use of inputs.

Non-linear models are presented in Table 2 (equations (5)-(6)), with technological progress proxied by a trend.[13] The estimated parameters for the input accumulation rates in the steady-state equations (column 1) are quite far from those usually obtained in the empirical literature, the low value of α being particularly remarkable. The effect of inflation is negative and significant. The coefficients for the convergence equation, when estimated independently (column 2), are more reasonable with an implicit rate of convergence of around 2.7%. The estimated values for α and γ are slightly smaller than those found in the literature. Again, the effect of inflation is negative and

dummies have been included (instead of the trend) to proxy the deterministic component of growth. The parameters of the model do not change significantly and these results are not reported to save space.

[13] Alternative characterisations of technological progress have been tried. First, substituting time dummies for the trend and, second, imposing a value for the rate of technological progress equal to 2%. The estimated parameters for the basic variables of the growth model and for the effect of inflation do not change significantly.

significant. Both equations (steady-state and convergence) have been estimated jointly imposing the theoretical restrictions across the coefficients α, γ and μ. These restrictions are not rejected (although in some cases they are accepted only marginally). The input shares (coefficients of the production function) obtained in the joint estimation are near to 0.1, 0.6 and 0.3 for physical capital, human capital and labour, respectively, and the implicit value for the convergence rate is 3.3%.

Table 2
Non-linear models (equation (5)-(6))

Dependent variable:	**(1)** y^*	**(2)** Δy	**(3)**^{j} $y^*, \Delta y$	**(4)** Δy	**(5)** Δy
Ω^s	-3.02 (3.28)	-	-2.98 (3.97)	-	-
Ω^c	-	-3.97 (2.88)	-1.97 (2.85)	-1.7 (1.48)	-0.81 (0.60)
α	0.08 (1.38)	0.27 (2.79)	0.10 (1.92)	0.23 (2.46)	0.34 (3.46)
γ	0.35 (6.23)	0.27 (3.03)	0.33 (6.42)	0.21 (2.36)	0.12 (1.24)
ϕ_{SS}	0.045 (2.94)	-	0.048 (3.25)	-	-
ϕ_c	-	-0.090 (2.55)	-0.071 (2.85)	-0.05 (1.73)	-0.03 (1.31)
ϕ'	-	-	-	-	0.003 (1.57)
μ	-0.011 (3.29)	-0.012 (2.16)	-0.012 (4.12)	-0.02 (4.39)	-0.014 (1.89)
λ	-	0.027	0.033	0.032	0.031
R^2_{ss}	0.52	-	0.51	-	-
R^2_c	-	0.38	0.33	0.41	0.34
σ_{ss}	0.265	-	0.262	-	-
σ_c	-	0.053	0.054	0.050	0.055

Notes: Estimation method: see notes in Table 1. *j*: joint estimation with cross-equation restrictions.

$\chi^2(\alpha)$: 4.04; $\chi^2(\gamma)$: 3.80; $\chi^2(\mu)$: 2.51; $\chi^2(\alpha,\gamma,\mu)$: 5.48.

The sensitivity of the results to the exclusion of some countries has also been tested. More precisely, when high-inflation countries, such as Turkey, Greece, Portugal or Spain, are omitted from the sample the results do not change significantly. However, this is not the case with Iceland. When Iceland is excluded from the sample (column 4) the effect of inflation on growth is almost twice as high as when it is included. This is not surprising since Iceland, being the country with the second highest average inflation within the OECD, is also a high-income fast-growing economy which may be generating a downward bias in the absolute value of the growth-inflation correlation.

The negative effect of inflation on per capita income appears as a robust result both in the steady-state and in the convergence equation. Although the negative influence of inflation on per capita income is solidly established, the effect on the sustainable growth rate is less clear. If the inflation rate is a determinant of steady-state per capita income (y^*), it will also appear in the convergence equation. But it is not clear if the effect found in this equation is an effect on the level or on the growth rate. To discriminate between these effects, an alternative specification is pursued in which A_t has been defined as in (4') and the convergence equation was estimated from the system (5) and (6) (column 5). This specification distinguishes between an inflation effect on the adjustment component of the convergence equation (μ) – i.e. the effect on the steady-state per capita income level

– and an effect on the trend component of the growth rate (ϕ'). Both are negative and significant when they are introduced individually, but when they are jointly included the effect on the trend component is non-significant and, in some cases, positive. This suggests that inflation impinges on the level of per capita income but not on the long-run rate of growth of the economy. The impact on the growth rate will thus be transitory (in the medium run), as long as real convergence is under way.[14]

Finally, a relevant issue to address is the linearity of the effect of inflation, in particular whether the marginal cost of inflation is constant or varies according to the level of inflation. Unlike most of the empirical work in this field which uses the Summers-Heston data set, our sample is confined to OECD countries, which share certain institutional features but have rather different inflation performances. Thus the OECD sample may be useful in analysing the sensitivity of the relationship between inflation and growth to the level of inflation.

The different perspectives adopted to analyse the linearity of the inflation effect have led to contradictory results. For instance, Barro (1995), estimating different coefficients for different levels of inflation, finds a greater effect of inflation on growth the greater the inflation level. Motley (1994), estimating the growth model for different sub-samples, concludes the opposite. In this paper, both approaches are used to analyse the marginal impact of changes in inflation on growth. First, in line with Barro (1995), the following specification, including different coefficients for three different inflation intervals, has been estimated:

$$y_{T+\tau} - y_T = \psi_0 + \psi_1 T + \psi_2 y_T + \psi_3 S^*_{TK} + \psi_4 S^*_{TH} + \psi_5 \log\left(n^*_T + \phi + \delta\right) + \psi_a \log\pi_a + \psi_b \log\pi_b + \psi_c \log\pi_c \quad (7)$$

where π_a is the level of inflation if it is lower than 6% and 0 otherwise; π_b is the level of inflation if it belongs to the interval 6-12% and 0 otherwise; and π_c is the level of inflation if it is greater than 12% and 0 otherwise.

The inflation coefficients obtained so far can be interpreted as semi-elasticities. For the purpose of comparing the inflation costs across groups with (by construction) very different average inflation rates we focus instead on elasticities; thus the inflation rate is specified in logs.[15] The estimated values for these elasticities are reported in part A of Table 3. The elasticity of income with respect to inflation is negative but not statistically different across the different average inflation levels. As an alternative approach, the homogeneity assumption may be relaxed by estimating the convergence equation for different sub-samples. This approach allows all the parameters – and not only the coefficient of inflation – to vary across sub-samples. The results are summarised in part B of Table 3. The effect of inflation is negative and significant for both low- (and very-low-) as well as for high- (and very-high-) inflation countries.

Thus, the two alternative approaches yield very much the same result: the elasticity of income with respect to inflation is the same regardless of the level of inflation. Since we deal with elasticities, this means that OECD countries have quite similar proportional output gains for a given proportional reduction of their inflation rate. If anything, this tells us that it pays more for a low-inflation country than a high-inflation country to reduce the inflation rate by a given amount. By the same token, it is rather more costly for a low-inflation country to concede an additional (and permanent) point of inflation than it is for a country with a higher starting rate.

To sum up, the analysis performed in this section, in accordance with other empirical work, supports the existence of an adverse influence of inflation on growth. As regards the size of this effect, if we take the coefficient in column 4 as a reliable estimate of the long-run effect of inflation upon growth, an increase in average inflation of one percentage point reduces, during rather long

---

[14] As suggested by the discussant, it is interesting to emphasise that although there is not a growth effect, the effect on the level occurs during a rather long period.

[15] Iceland is excluded from the sample.

Table 3

Linearity of the inflation effect

Elasticities of income with respect to inflation in estimates of the linear version of the convergence equation

A. Whole sample estimates with specific inflation coefficients	
Low inflation[1]	-0.101 (2.24)
Medium inflation[1]	-0.066 (2.44)
High inflation[1]	-0.064 (4.08)
B. Sub-sample estimates	
Low inflation[2]	-0.045 (2.51)
High inflation[2]	-0.037 (2.41)
Very low inflation[3]	-0.040 (1.93)
Very high inflation[3]	-0.052 (2.21)

Notes: Estimation method: see notes in Table 1. Absolute *t*-ratios in parentheses.

[1] Low inflation: observations with inflation lower than 6%; Medium inflation: observations with inflation between 6% and 12%; High inflation: observations with inflation greater than 12%.

[2] Low inflation: countries with average inflation lower than the median; High inflation: countries with average inflation greater than the median.

[3] Very low inflation: 8 countries with the lowest inflation. Very high inflation: 8 countries with the highest inflation.

periods, per capita growth by 0.08% per year, which leads to a reduction in steady-state per capita income of about 2 percentage points. In the next two sections we take a closer look at the relationship between inflation and growth by trying more general specifications to avoid the two main drawbacks of the methodology adopted so far: the assumption of homogeneity across countries of the technological parameters, and the direction of causality among the variables in the convergence equations.

3. Country-specific effects and the cost of inflation

There are several reasons for including individual effects in convergence equations estimated with multi-country data sets. The existence of technological differences in the rates of technical progress or, as is more likely, in the initial conditions of each country, will be reflected in the relevance of idiosyncratic effects in growth equations. Empirical analysis of the characteristics of economic growth has relied quite often on the use of information for wide groups of countries. It is extremely important to exploit the cross-section variability of multi-country data sets, given the way in which long-term growth and shorter-term fluctuations interact, and use time series averages. However, this approach imposes the very strong constraint that the data for all the economies of the sample stem from the same theoretical distribution; i.e. technological parameters are homogeneous across countries. Usually, this assumption is not explicitly tested, although its empirical implications may be very important.[16] The consideration of individual effects in the constant term (Knight et al. (1993), Islam (1995)) or, in a more general way, (Andrés, Boscá and Doménech (1996)) may alter significantly some of the main empirical results.

Table 4

Convergence equation with individual effects (equation (5'))
Dependent variable: Δy

	(1)	(2)	(3)	(4)	(5)
ψ	1.41 (2.00)	1.18 (8.18)	1.14 (2.77)	0.93 (3.07)	1.06 (10.15)
ψ_1	0.03 (3.48)	0.03 (4.28)	0.03 (3.76)	0.02 (3.63)	0.02 (4.09)
ψ_2	-0.55 (6.12)	-0.49 (7.95)	-0.46 (6.80)	-0.44 (7.38)	-0.42 (8.06)
ψ_3	-0.11 (1.32)	-	-0.02 (0.55)	-0.16 (0.67)	-
ψ_4	-0.001 (0.13)	-	-0.001 (0.03)	-0.04 (1.40)	-
ψ_5	-0.12 (0.64)	-	-0.02 (0.20)	-0.02 (0.26)	-
ψ_6	-0.001 (1.33)	-0.002 (2.32)	-0.002 (2.48)	-0.002 (2.82)	-0.003 (3.21)
R^2	0.54	0.59	0.57	0.55	0.55
σ	0.050	0.047	0.047	0.047	0.046

Notes: Estimation method: instrumental variable. Instruments: constant, trend, country dummies, initial per capita income, accumulation rate of human capital and first and second order lags of initial per capita income, inflation, inflation variability, accumulation rates of human capital, investment rate and growth rate of population.

Dummy variables included:
Columns (1) and (2): one for each country, except Australia.
Column (3): one for each of the following countries: Canada, Switzerland, Germany, Spain, Finland, the United Kingdom, Greece, Ireland, Iceland, Luxembourg, New Zealand, Portugal, Turkey and the United States.
Columns (4) and (5), one for each of the following countries: Spain, Greece, Iceland and Turkey and one for each of the following country groups: Ireland and Portugal; Canada and Germany; Switzerland, Luxembourg and the United States; and Finland, New Zealand and the United Kingdom.

In this section, by adding individual effects to the linear version of the convergence equation in (5'), we test if the negative effect of inflation on growth is due to a misspecification caused by the omission of individual effects. The main results are summarised in Table 4. Including individual effects in the convergence equation does not involve additional problems in the estimation process, given that the availability of eight time observations allows a specific constant to be estimated for every country. All the models have been estimated by instrumental variables in order to correct the simultaneity bias caused by the endogeneity of some of the regressors (inflation in particular).[17] When we add a dummy variable for each country (column 1) the explanatory power of most regressors changes, as compared with the models in the previous section. In particular, while inflation still has a negative effect on income, the t-statistic of the rate of inflation is now lower (1.33). The changes in the rest of the model are far more radical. First, whereas a negative trend

[16] See Pesaran and Smith (1995).

[17] Instruments are listed in Table 4.

coefficient was an unappealing feature of the models in Section 2, this coefficient now becomes positive and significant, with a reasonable point estimate of 0.03. Secondly, the point estimates of the technological coefficients are now either non-significant or wrongly signed. In fact, when excluding the accumulation rates from the equations, the negative correlation between growth and inflation becomes highly significant with a t-statistic of -2.32 (column 2). Finally, several country dummies are not different from zero, suggesting that estimating a more parsimonious model is advisable.

Starting from the model with a dummy variable for each country, the non-significant dummy variables have been removed, setting aside the one with the lowest t-statistic each time. As a second step, these excluded variables have been added again, one at a time, retaining those with a t-ratio greater than 1.5. Every time a dummy variable is added back into the model, the process is reinitiated. This procedure does not involve the analysis of every single possible specification according to all the combinations of country-specific constants. However, it provides a model selection procedure that allows us to test, at least twice, the marginal significance of each dummy variable: first, against a more general model (with all the country-specific dummies); and next, against a more restricted one. This procedure allows the contribution of each dummy variable to be assessed, not only in terms of fitting but also in terms of its influence on the coefficients of the rest of the variables.

The estimate in column 3 summarises the final outcome of this specification process. The results do not change very much with respect to those in column 1 and the coefficient of the inflation rate remains negative and significant, with a size similar to that obtained for the model without individual effects. Furthermore, this result is quite robust to the set of country-specific dummies included in the regression. As in column 1, the model does not reproduce other results of the exogenous growth model: more precisely, the coefficients for the investment rate and for the accumulation rate of human capital are still non-significant.

Taking column 3 as a starting point, the model in column 4 substitutes country-group dummy variables for the individual country dummy variables. The country groups have been defined according to the size of the individual effect. Greece shows an individual effect that is clearly negative (-0.31) as compared with the excluded countries,[18] followed by Turkey (-0.29), Ireland and Portugal (-0.22), Spain (-0.15), and New Zealand, Finland and the United Kingdom (-0.05). On the other hand, Canada and Germany (0.04), Iceland (0.08) and Switzerland, Luxembourg and the United States (0.1) display positive individual effects on the growth rate. The estimated individual effects reveal a systematic pattern which, if ignored, could bias the estimated effect of inflation. The individual effect is strongly correlated with the level of per capita income achieved at the end of the sample period. Thus, omitting the individual effect, the model would underestimate the growth of the richest countries, while overestimating that of the poorest countries. Since there is a negative correlation, at the OECD level, between per capita income in 1993 and the average inflation rate, excluding the individual effects is a source of potential bias in the estimation of the effect of inflation.

In column 4, the t-statistic of the inflation rate has increased again (up to -2.82); however, as in column 1, the coefficients for the input accumulation rates are not significant. In fact, the exclusion of these variables (column 5) does not worsen substantially the fit of the equation and further increases the significance level of the inflation rate. It is quite remarkable that the negative and highly significant influence of inflation on growth during rather long periods survives all these changes in the specification. In fact, together with the initial per capita income, it turns out to be the most robust variable of the model. As for the size of the effect of inflation on per capita income in the long-run, it is also similar to that obtained in the models without individual effects. Thus, according to columns 3 and 4, a reduction of one percentage points in the average inflation rate would increase real per capita income of around 0.06 percentage points per year, which leads to a permanent rise in the

---

[18] Australia, Austria, Belgium, Denmark, France, the Netherlands, Italy, Japan, Norway and Sweden.

steady-state per capita income of about 2%. The size of this effect is smaller than that obtained in the models without individual effects.[19]

4. Inflation and growth: analysis of causality

The models studied in previous sections can give rise to a notable bias in the estimation of the influence of inflation on growth by focusing on the contemporaneous correlation between these two variables. Inflation and growth are the joint outcome of the way in which an economy responds to different shocks. If demand shocks predominate, a positive association between GDP growth and inflation can be expected, whereas the association will be negative in response to supply shocks. Also, even if we consider the possibility of a true influence of one variable over the other, the theoretical literature presents arguments in favour of causality in either of the two directions. For this reason, the contemporaneous correlation between growth and inflation may not be very informative as to the existence and magnitude of a real cost associated with inflation.

This section analyses the statistical causality, as formulated by Granger, between inflation and growth. This perspective is broader than that of convergence equations in several ways. First, the analysis of causality focuses on the study of non-contemporaneous effects and the marginal explanatory power of one variable over the other. This is precisely the influence of inflation on growth assigned by the theoretical models: an influence that does not operate in the short run but instead manifests itself slowly through a reduction in factor accumulation or in the deterioration of its efficiency. Second, in using a more flexible specification, we avoid the imposition of the parametric restrictions that are often included in the neo-classical growth model and which can affect the test of the correlation that concerns us here. The analysis of causality can, however, benefit from the teachings of the theoretical framework for the convergence equations by suggesting a series of determinants of growth that can be incorporated into the information set in the causality tests.

4.1 The econometric model

The Granger test of causality is a test of exclusion in the following bivariate VAR system:

$$\begin{bmatrix} \Delta y_t \\ \pi_t \end{bmatrix} = \begin{bmatrix} A \\ B \end{bmatrix} + \begin{bmatrix} C(L) & D(L) \\ E(L) & F(L) \end{bmatrix} \begin{bmatrix} \Delta y_t \\ \pi_t \end{bmatrix} + \begin{bmatrix} u_t \\ \varepsilon_t \end{bmatrix} \tag{8}$$

where y_t and π_t are vectors (24x1) of current observations of the logarithm of per capita GDP and of the rate of inflation, respectively, for the 24 member countries of the OECD, and A and B are vectors (24x1) of constants. $C(L)$, $D(L)$, $E(L)$ and $F(L)$ are matrices of order (24x24) in which the elements outside the main diagonal are zero and the elements on the main diagonal form a lagged polynomial of order p such that for $C(L)$:

$$c_1 L + c_2 L^2 + c_3 L^3 + \ldots + c_p L^p$$

The elements of the matrices A and B, as well as the coefficients of the lagged polynomials (in $C(L)$, $D(L)$, $E(L)$ and $F(L)$), will be assumed to be homogeneous among countries unless expressly stated otherwise.

---

[19] We also performed several tests on the linearity of inflation elasticity in models with individual country dummies. The results, not reported here to save space, display a similar pattern to those in Table 3. The coefficient of inflation remains significant in all cases, with no significant differences across inflation levels.

The rejection of the null hypothesis that the d_j are zero indicates that current inflation helps to reduce the mean-squared error in the prediction of per capita income and, therefore, that π causes Δy as formulated by Granger. Likewise, the rejection of the null hypothesis that the coefficients e_j are zero indicates that Δy causes π. On the other hand, testing the causality between y_t and π_t only entails eliminating a linear restriction on the coefficients in $C(L)$ and $E(L)$, such that per capita income cannot be written in first differences. The estimation of equation (8) gives rise to several methodological questions concerning the information set, the presence of individual effects, the appropriateness of subjecting the tests to a broader information set, and the non-stationarity of the model's variables. Since this section applies annual data relating to the variables of interest for the 24 OECD countries, it departs somewhat from the traditional approach in the empirical literature on growth, which avoids using annual information by applying more or less refined filtering procedures.[20] As to the possibility of individual effects relating to each OECD country, we noted in the previous section that their presence may be necessary to correct the biases in the estimated correlation between π and y. We will include these effects in this section by considering several specifications in which the vectors A and B include a different constant for each country (a_i, b_i).

The most important problem posed by the exclusion test in equation (8) is the possibility that some variables are not stationary, in which case these tests do not have a standard distribution. In the case at hand, both per capita income and the rate of inflation are non-stationary.[21] In the literature, several different methods are proposed for testing the hypothesis of causality between integrated variables, making use of statistics with asymptotic distribution $\chi^2(p)$. These procedures basically consist of re-parametrising the model in order to obtain stationary regressors.[22] Of these approaches, the method proposed by Dolado and Lütkepohl (1994) does not require a search for possible cointegration vectors among the model's variables. They propose the estimation of a VAR in levels of order $p+1$. The exclusion test performed on the p first lags[23] is thus distributed asymptotically as an F, whereby the loss of efficiency by the over-parametrisation of the model is compensated by the test's consistency and simplicity.[24] The application of this method requires knowing the true order, p, of the VAR. In this paper, rather than discussing the structure of the lags in detail, we will present results for a sufficiently broad range of lags which ensure that the residuals are stationary.

4.2 Results

To analyse the causality from the rate of inflation to the growth rate of per capita income, a test is run on the joint significance of $\{d_1,...,d_p\}$ in the model:

$$\Delta y_t = A + C(L)\Delta y_t + D(L)\pi_t + G(L)X_t + u_t \tag{9}$$

where $G(L)$ is a matrix of a structure similar to $C(L)$ and X_t is a vector of additional regressors, suggested by economic growth theory and whose composition is defined for each specification. Likewise, the causality from the growth rate to inflation is tested through the joint significance of $\{e_1,...,e_p\}$ in:

---

[20] An increasing number of studies tend to use raw annual data. Moreover, in the dynamic analysis of causality, models based on averages (five-year averages, for example) can be considered restricted versions of models that use annual data.

[21] At least for many of the countries in the sample.

[22] See Sims, Stock and Watson (1990).

[23] In other words, on d_1, d_2,..., d_p or on e_1, e_2,..., e_p.

[24] For an application of this method, see Andrés, Boscá and Doménech (1996).

$$\pi_t = B + E(L)\Delta y_t + F(L)\pi_t + H(L)X_t + \varepsilon_t \qquad (10)$$

Also, to analyse causality from the level of per capita income, a test is performed on the significance of the parameters $\{e_j\}$ and $\{d_j\}$ in (10'), obtained by replacing the growth rate of per capita income (Δy_t) in (10) by its level (y_t).

The first of these tests (Table 5) summarises the significance level for the null hypothesis that the inflation coefficients are jointly non-significant in diverse models. In addition, it gives the t statistic for the sum of these coefficients. All the models were estimated for a structure of lags (p+1) that goes from 4 to 13, and the related tests were run for the p first lags in each case. In the first column, the model estimated (model 1) includes no additional regressor (g_j = 0) nor constant individual effects (A is a vector of equal constants). The results of this first exercise are not very conclusive. In most of the specifications, inflation appears to cause the growth rate as formulated by Granger (the significance level of the statistic only marginally exceeds 5% for p=8 and for p=10). However, the sign of this causality is far from unequivocal, because in general the sum of the coefficients of the inflation lags is not significantly different from zero. In model 2, individual constant effects were included such that vector A is a vector of different constants by country. Although in this case causality can be rejected at 5%, for long inflation lags, the influence of inflation becomes more sharply negative. The t statistic is negative in most of the specifications and clearly significant for recent inflation (p=3).

Table 5

Causality from inflation to growth

p	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
	CHI	T	CHI	T	CHI	T	CHI	T	CHI	T
3	0.00	-1.57	0.00	-2.26	0.00	-2.73	0.02	-2.05	0.00	-2.97
4	0.00	-0.63	0.00	-1.25	0.00	-1.45	0.02	-0.68	0.00	-1.85
5	0.00	-0.29	0.00	-1.14	0.00	-1.71	0.01	-0.67	0.00	-1.67
6	0.02	0.52	0.02	-0.38	0.00	-0.92	0.03	-0.37	0.01	-1.26
7	0.04	0.18	0.06	-0.78	0.00	-1.58	0.05	-0.80	0.01	-1.11
8	0.06	0.59	0.10	-0.48	0.02	-1.56	0.08	-0.76	0.03	-1.45
9	0.02	-0.75	0.05	-1.51	0.00	-2.54	0.02	-1.81	0.01	-2.31
10	0.06	-1.29	0.14	0.34	0.02	-0.91	0.04	-0.45	0.02	-0.91
11	0.00	-0.64	0.05	-1.20	0.00	-2.52	0.01	-1.87	0.01	-2.06
12	0.02	0.00	0.17	-0.41	0.01	-1.87	0.13	-1.35	0.11	-1.41

Notes: Model 1: Basic model with no individual effects nor additional regressors. Model 2: Model with individual effects. Model 3: Model with individual effects and accumulation rates (investment rate, schooling, rate of growth of population and trend). Model 4: Model with individual effects and lagged (t-1) accumulation rates. Model 5: As in model 4 including some lagged (t-1) macroeconomic indicators (money growth, exports growth and public consumption as a percentage of GDP).

Models 3 and 4 incorporate, in addition to individual effects, several other regressors such as a linear trend, the savings ratio, the rate of schooling and the growth rate of the population. In model 3 these regressors are included contemporaneously, whereas their first lags are included in model 4 to avoid biases of simultaneity. The incorporation of these regressors leads to a substantial fall in the value of the F, with an increase in the absolute value of the t statistic. In model 3 the null hypothesis of non-causality is rejected for any value of p; moreover, the sign of the long-run influence of inflation on the growth rate is negative in all cases, although it is not always significant at the 5 per cent level. Model 4 displays a very similar pattern to that of model 3, although both the causality and the negative sign of the influence of inflation on growth are less clear-cut.

Model 5 incorporates, in addition to the regressors of model 4, the first lags of several other macroeconomic variables such as the growth rate of the money supply, the growth rate of exports and the growth of public spending as a percentage of GDP. Many authors have studied the relationship between long-term growth and the short-term performance of economies.[25] The main argument on which this relationship rests is that the shocks in an economy or the way economic policy is conducted influence agents' accumulation decisions and the way markets operate. Thus, a succession of negative shocks or an inadequately designed fiscal or monetary policy may have effects that go beyond the short term, affecting potential output and sustained growth. If this argument is correct, the causal interpretation of the estimated correlation between inflation and growth could be called into question as it could be due to the fact that inflation approximates the impact of other macroeconomic variables with which it is, in turn, strongly correlated.

A specification that includes other macroeconomic indicators has the advantage that it permits an analysis of the influence of inflation on growth while isolating it from the effect of other shocks. The variables chosen have often been considered in empirical models of economic growth, albeit with different results as to their influence in the long run.[26] After taking into account the effect of fiscal and monetary policy and the performance of competitiveness, the existence of causality and a negative sign from inflation to economic growth become more apparent. Once again the null hypothesis that inflation does not help to improve the prediction of the future growth rate is clearly rejected at 5% and even at 1% (except for p=12). The statistic associated with the sum of the coefficients of the inflation lags in (8) is negative in all cases, with a value of t higher than unity (except for p=10), and significantly different from zero (at the 5% significance level) in four cases.

Table 6
Causality from inflation to per capita income

p	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
	CHI	T	CHI	T	CHI	T	CHI	T	CHI	T
3	0.00	-3.37	0.00	-2.85	0.00	-3.03	0.01	-2.24	0.00	-3.18
4	0.00	-1.84	0.00	-1.25	0.00	-1.52	0.04	-0.43	0.01	-1.67
5	0.00	-2.57	0.00	-1.83	0.00	-2.27	0.02	-1.12	0.00	-2.18
6	0.00	-1.73	0.00	-1.30	0.00	-1.18	0.02	-0.56	0.00	-1.57
7	0.00	-1.52	0.06	-1.10	0.01	-1.28	0.06	-0.52	0.01	-0.94
8	0.01	-1.46	0.01	-1.24	0.03	-1.55	0.09	-0.73	0.02	-1.38
9	0.01	-1.88	0.01	-1.80	0.01	-2.20	0.05	-1.39	0.02	-2.04
10	0.00	-1.07	0.01	-1.40	0.01	-1.40	0.03	-0.94	0.01	-1.36
11	0.01	-1.47	0.02	-1.65	0.01	-1.89	0.03	-1.38	0.01	-1.78
12	0.00	-1.54	0.02	-2.10	0.00	-2.25	0.01	-1.88	0.02	-2.03

Notes: see Table 5.

Table 6 shows the statistics of joint significance and of the sum of the parameters $\{d_j\}$ in specifications similar to those of Table 5, but without imposing the restriction of first differences in the dynamics of per capita income.[27] In most of these models the non-causality of inflation to per

[25] See Fischer (1993), for example.

[26] See Levine and Renelt (1992) and Andrés, Doménech and Molinas (1996a) for alternative views of the influence of macroeconomic shocks on growth.

[27] Causality analysis cannot discriminate between the *level* and *growth rate* effects, since the model in levels is merely a less restricted version of the model specified in differences. The purpose of this exercise is simply to check whether the results given in Table 5 depend on the dynamic specification.

capita income is rejected, even at the 1% significance level. Moreover, the sign of the influence of inflation on income is always negative, with a t statistic often higher than 1.5 and significant (at 5%) in a third of the specifications applied. The inclusion of individual effects (model 2 versus model 1) marginally alters the estimated effect of inflation on per capita income. The inclusion of additional regressors has a similar effect on the statistics. Model 4 reflects an across-the-board fall in the joint significance of the inflation lags, not only due to the increase in the significance level of the F (which moves close to or higher than 5%, but in no case reaches 10%) but also due to the decline in the absolute value of the t of the sum of the coefficients of the lags. This variation (also apparent in the analysis of causality from π_t to Δy_t) could be interpreted in terms of the channel through which the influence of inflation manifests itself. In particular, the increase in the significance level at which the null of non-causality is rejected and the decrease in the size of the long-run coefficient of inflation would indicate that causality is largely explained by the investment channel. Finally, once the effect of other short-run macroeconomic indicators is taken into account, the effect of inflation on income is seen more clearly. In model 5, once again, the value of the F is high enough to reject the null hypothesis of non-causality in all cases (at the 5% level). Moreover, the sum of the inflation lags is significantly negative in four cases at the 5% level and at slightly lower significance levels in three other cases. We have finally tested the existence of reverse causality from the rate of economic growth onto the rate of inflation. As can be seen in Table 7, causality is only rejected in one out the 50 specifications presented. Thus we may conclude that current growth rates help to explain the future course of the inflation rate. The t statistic of the long-run coefficient is always positive and significant in more than 90% of the cases. The size of this effect peaks at p equal to 4 falling thereafter. The inflationary effects of growth are subject to a certain delay and, as expected, they diminish over time.

Table 7
Causality from growth to inflation

p	**Model 1**		**Model 2**		**Model 3**		**Model 4**		**Model 5**	
	CHI	T	CHI	T	CHI	T	CHI	T	CHI	T
3	0.00	3.12	0.00	3.18	0.01	2.62	0.02	2.42	0.14	1.89
4	0.00	5.76	0.00	5.66	0.00	4.51	0.00	4.48	0.00	4.15
5	0.00	4.99	0.00	4.72	0.00	3.88	0.00	3.34	0.00	2.98
6	0.02	5.15	0.00	4.69	0.00	3.91	0.00	3.53	0.00	3.40
7	0.00	4.38	0.00	4.38	0.00	3.14	0.00	2.55	0.00	2.24
8	0.00	4.17	0.00	4.00	0.00	3.03	0.00	2.49	0.00	2.51
9	0.00	3.76	0.00	3.72	0.00	2.74	0.00	2.19	0.00	2.38
10	0.00	3.86	0.00	3.87	0.00	2.61	0.00	1.94	0.00	2.18
11	0.00	2.23	0.00	3.03	0.00	1.77	0.00	1.04	0.00	1.37
12	0.00	2.09	0.00	2.58	0.01	0.87	0.01	0.28	0.03	0.63

Notes: see Table 5.

Economic theory proposes several explanations why rapid growth is associated with higher near-term inflation. It could be a movement along a negatively sloped Phillips curve, as prices respond after a period of rapid expansion in demand. Another interpretation is derived from the so-called Balassa-Samuelson effect,[28] whereby rapid economic growth is associated with rapid expansion in the productivity of a country's tradable goods sector, in turn leading to an appreciation of its currency. Insofar as the nominal exchange rate is not adjusted to produce this appreciation, domestic prices will grow faster. This positive correlation indicates that the risk of a simultaneity bias in the estimation of inflation costs is considerable. However, this effect creates a downward bias in the

[28] Balassa (1964) and Samuelson (1964).

estimated impact of inflation on growth.[29] As a result, the contemporaneous correlation in the convergence equations could be regarded as a lower bound of the costs of inflation which would have to be adjusted upwards.

The statistics summarised in Tables 5, 6 and 7 allow us to draw a first general conclusion as regards the importance of the effects studied here. Current inflation helps forecast future income. Although the negative effect of inflation on per capita income in the long run is not always significant, the sign of the sum of the lagged inflation coefficients in models (10) and (10') is never positive or significantly higher than zero. Thus, a high level of inflation today may or may not be bad news but it can never be taken as good news, as far as the future income prospects of an economy are concerned.

The importance of these results becomes more apparent if the causality from inflation to income is compared with the influence on growth of other variables such as investment in physical or human capital or the expansion in public spending. Several recent attempts to corroborate the statistical causality from investment in physical capital to growth and income have concluded that, even though it cannot be rejected that a high rate of current investment could be explained by rapid growth in the past, the existence of causality in the opposite direction is far less conclusive. Blömstrom, Lipsey and Zejan (1996) show, for a group of 101 countries, how growth always precedes investment in simple growth equations. A similar result is obtained by Carrol and Weil (1993) for the OECD sample. Andrés, Boscá and Doménech (1996) also find that investment does not help to improve the prediction of income or of its growth rate in practically any of the specifications studied. Moreover, when investment appears to "cause" income, the negative sign makes this result hard to interpret.[30] A similar effect is obtained in relation to the other fundamental variable in the determinants of growth, the rate of schooling, as a proxy for investment in human capital. The authors argue that the estimated correlation in most growth equations between income and investment can be explained by the existence of a simultaneity bias or by the exclusion of country-specific effects in these equations. The existence of a simultaneity bias cannot be rejected because, according to the above-mentioned studies, income helps to explain investment, in the sense that the higher the income, the higher the rate of saving and investment of OECD economies.

The results presented in this section have an unequivocal interpretation. The current rate of inflation provides relevant information on income prospects in OECD countries. In particular, ceteris paribus, higher inflation anticipates a lower level of income in the medium and long run. This effect is robust to alternative specifications and, most notably, survives even when accumulation rates and individual effects are included among the set of regressors. This would indicate that inflation can influence not only saving and investment in economies but also the efficient use of productive factors. Moreover, it can be rejected that this leading correlation between inflation and income is spurious and produced by the coincidence of inflationary tendencies and slow growth in some economies. Therefore, even though the magnitude of the negative effect of inflation might be questioned, the results of this section allow us to affirm that inflation does not appear to be neutral in the long run and that in no case does the persistence of inflationary tensions favour rapid economic growth in the future.

Concluding remarks

The purpose of this paper has been to study the correlation between growth and inflation at the OECD level, within the framework of the so called convergence equations. It further discusses

---

29 In this respect, see Andrés, Hernando and Krüger (1996).

30 This is consistent with the findings by Carrol and Weil (1993).

whether this correlation withstands a number of improvements in the empirical models, which try to address the most common criticisms on this evidence. Despite its many shortcomings, this approach is well-suited to test the robustness of the correlation between growth and inflation in low-inflation economies with reasonably well-working markets, such as the OECD countries during the 1960-92 period.

The main finding is that there is a significant negative correlation between inflation and income growth during rather long periods. This negative correlation survives the presence of additional regressors, such as savings, population growth and schooling rates, and the imposition of the theoretical restrictions implied by the constant returns of technology. What is most remarkable is that the negative coefficient of inflation in growth equations remains significant even after allowing for country-specific time-invariant effects in the equations. This is striking since, as is well known in the empirical growth literature, few regressors in convergence equations withstand the explanatory power of country dummies. The analysis of causality gives less clear-cut results, but it is also noteworthy that causality from inflation to growth is always significant and never positive. Again, this result shows up most clearly when the influence of country dummies, accumulation rates and the effect of other macroeconomic variables is controlled for.

According to the point estimate of its coefficient, inflation has a negative but temporary impact upon long-term growth rates, which, in turn, generates a permanently lower level of per capita income than would otherwise have been achieved. The estimated cost of an additional 1% in the inflation rate is a reduction, during rather long periods, in the annual growth rate of about 0.06%, that leads to a reduction in the steady-state per capita income of about 2%. Inflation not only reduces the level of investment but also the efficiency with which productive factors are used. Our results seem to suggest that the marginal cost of inflation is independent of the inflation rate. Nonetheless, the long-run costs of inflation are non-negligible and efforts to keep inflation under control will sooner or later pay off in terms of better long-run performance and higher per capita income.

References

Andrés, J., E. Boscá and R. Doménech (1996): "Testing the neoclassical growth model: a causality approach". *Mimeo.*

Andrés, J., R. Doménech and C. Molinas (1996a): "Growth and Convergence in OECD Countries: A Closer Look", in B. van Ark and N. Crafts (eds.), *Catch Up and Convergence in Post War Europe: Quantitative Aspects*, Cambridge University Press.

Andrés, J., R. Doménech and C. Molinas (1996b): "Macroeconomic Performance and Convergence in OECD Countries". *European Economic Review* (forthcoming).

Andrés, J., I. Hernando and M. Krüger (1996): "Growth, inflation and the exchange rate regime". *Economics Letters*, No. 53, pp. 61-65.

Balassa, B. (1964): "The purchasing power parity doctrine: a reappraisal". *Journal of Political Economy*, No. 72, pp. 584-596.

Barro, R. and X. Sala-i-Martín (1991): "Convergence Across States and Regions". *Brookings Papers on Economic Activity*, No. 1.

Barro, R. (1995): "Inflation and economic growth". *Bank of England Economic Bulletin*, pp. 1-11.

Blömstrom, M., R. Lipsey and M. Zejan (1996): "Is fixed investment the key to economic growth?" *Quarterly Journal of Economics*, No. 111, pp. 269-276.

Briault, C. (1995): "The costs of inflation". *Bank of England Economic Bulletin*, pp. 33-45.

Bruno, M. (1993): "Inflation and Growth in an integrated approach". *NBER Working Paper*, No. 4422.

Burdekin, R., T. Goodwin, S. Salamun and T. Willett (1994): "The effects of inflation on economic growth in industrial and developing countries: is there a difference?" *Applied Economic Letters*, No. 1, pp. 175-177.

Cardoso, E. and A. Fishlow (1989): "Latin American Economic Development: 1950-1980". *NBER Working Paper*, No. 3161.

Carrol, C. and D. Weil (1993): "Saving and growth: a reinterpretation". *NBER Working Paper*, No. 4470.

Clark, T.E. (1993): "Cross-Country Evidence on Long Run Growth and Inflation". *RWP 93-05*, Federal Reserve Bank of Kansas City.

Cozier, B. and J. Selody (1992): "Inflation and Macroeconomic Performance: Some Cross-Country Evidence". Department of Monetary and Financial Analysis, Bank of Canada.

Dabán, T., R. Doménech and C. Molinas (1996): "International and intertemporal comparisons in OCDE countries: a growth sensitivity analysis". *Review of Income and Wealth* (forthcoming).

De Gregorio, J. (1992a): "Economic Growth in Latin America". *Journal of Development Economics*, No. 39, pp. 59-84.

De Gregorio, J. (1992b): "The effects of inflation on economic growth: Lessons from Latin America". *European Economic Review*, No. 36, pp. 417-425.

De Gregorio, J. (1993): "Inflation, Taxation, and Long-run Growth". *Journal of Monetary Economics*, No. 31, pp. 271-298.

De Gregorio, J. (1994): "Inflation, Growth and Central banks: Theory and Evidence". *Mimeo.*

Dolado, J. and H. Lütkepohl (1994): "Making Wald test work for cointegrated VAR systems". *CEMFI Working Paper*, No. 9424.

Feldstein, M. (1996): "The costs and benefits of going from low inflation to price stability". *NBER Working Paper*, No. 5469.

Fischer, S. (1991): "Growth, Macroeconomics, and Development". *NBER Working Paper*, No. 3702.

Fischer, S. (1993): "The Role of Macroeconomic Factors in Growth". *Journal of Monetary Economics*, No. 32, pp. 485-512.

Grier, K. and G. Tullock (1989): "An empirical Analysis of Cross-National Economic Growth, 1951-80". *Journal of Monetary Economics*, No. 24(2), pp. 259-276.

Grimes, A. (1991): "The Effects of Inflation on Growth: Some International Evidence". *Weltwirtschaftliches Archiv*, No. 127(4), pp. 631-644.

Islam, N. (1995): "Growth empirics: a panel data approach". *Quarterly Journal of Economics*, No. 110, pp. 1127-1170.

Jones, L. and R. Manuelli (1993): "Growth and the effects of inflation". *NBER Working Paper*, No. 4523.

Knight, M., N. Loayza and D. Villanueva (1993): "Testing the neoclassical theory of economic growth: a panel data approach". *IMF Staff Papers*, No. 40, pp. 512-541.

Kormendi, R., and P. Meguire (1985): "Macroeconomic Determinants of Growth. Cross-Country Evidence". *Journal of Monetary Economics*, No. 16(2), pp. 141-163.

Levine, R. and D. Renelt (1992): "A Sensitivity Analysis of Cross-Country Growth regressions". *American Economic Review*, No. 82, pp. 942-963.

Levine, R. and S. Zervos (1993): "What Have We Learned about Policy and Growth from Cross-Country Regressions? *American Economic Review*, No. 83, pp. 426-430.

Mankiw, N., D. Romer and D. Weil (1992): "A Contribution to the Empirics of Economic Growth". *Quarterly Journal of Economics*, No. 107, pp. 407-38, May.

Motley, B. (1994): "Growth and Inflation: A Cross-Country Study". Federal Reserve Bank of San Francisco, *Working Paper*, No. 94-08.

Orphanides, A. and R. Solow (1990): "Money, inflation and growth", in B. M. Friedman and F. H. Hahn (eds.), *Handbook of Monetary Economics*, Vol. 1.

Pesaran, M. H. and R. Smith (1995): "Estimating long-run relationships from dynamic heterogeneous panels". *Journal of Econometrics*, No. 68, pp. 79-113.

Pindyck, R. and A. Solimano (1993): "Economic instability and aggregate investment" *NBER Working Paper*, No. 4380.

Roubini, N. and X. Sala-i-Martin (1995): "A growth model of inflation, tax evasion and financial repression". *Journal of Monetary Economics*, No. 35, pp. 275-301.

Sala-i-Martin, X. (1994): "Cross-sectional Regressions and the Empirics of Economic Growth". *European Economic Review*, No. 38, pp. 739-747.

Samuelson, P. (1964): "Theoretical notes on trade problems". *Review of Economics and Statistics*, No. 46, pp. 145-154.

Sims, C., J. Stock and M. Watson (1990): "Inference in linear time series models with some unit roots". *Econometrica*, No. 58, pp. 113-44.

Smyth, D. (1994): "Inflation and Growth". *Journal of Macroeconomics*, Vol. 16, pp. 261-270, spring.

Solow, R. (1956): "A Contribution to the Theory of Economic Growth". *Quarterly Journal of Economics*, No. 70, pp. 65-94.

Comments on: "Inflation and economic growth: some evidence for the OECD countries" by Javier Andrés and Ignacio Hernando

by R. Sean Craig

This paper is an interesting contribution to the cross-country growth regression literature that tests for an effect of inflation on economic growth. As is well known, this literature is plagued by two problems:

1) estimation results are not robust to changes in specification or the addition of variables;
2) Problems of interpretation arising from the fact that the direction of causality may be the reverse of that assumed in the model.

The paper attempts to address these problems using panel data to estimate the relationship between growth and inflation in OECD countries. It does this in three ways: first, rather than estimate a reduced-form regression, as is often done in the literature, it estimates a structural model based on a fairly sophisticated version of the Solow growth model to help clarify the problem of interpretation noted above. Second, the paper subjects the model to a number of robustness tests. Third, it uses Granger causality tests to test whether the causality does in fact run from inflation to growth. These latter tests generally show that the causality runs in the appropriate direction, and represent a significant contribution to the literature.

The structural model allows for a number of different variables to influence growth: inflation, physical and human capital accumulation, population, the level of income and a time trend. In the model, inflation lowers per-capita income by reducing the efficiency of investment. While the use of a theoretical model is a nice feature of this paper, its usefulness is limited by the fact that it does not allow for other channels through which inflation could reduce growth that could be important. Specifically, it does not incorporate the effect, identified by Martin Feldstein, where, through the interaction of inflation and the tax system, inflation increases the after-tax cost of capital. Nor does it allow for the possibility that inflation, or inflation variability, can lower the investment rate by distorting inter-temporal decision-making.

The paper does not find a relationship between inflation and the *growth rate* of GDP but, rather, a statistically significant negative relationship between inflation and the *level* of per-capita GDP. However, since this latter effect occurs over 30 years it is quantitatively very similar to a growth-rate effect. Moreover, in estimation it is hard to distinguish these two effects statistically. Also, the paper does not find an effect of the variability of inflation on growth or per-capita income. This may reflect the limitations of the data since the level of inflation and the variability of inflation are highly correlated so it would be difficult to identify separate effects from these two variables.

While the negative relationship between inflation and per capita income remains statistically significant for all specifications that are estimated, the magnitude of the effect is sensitive to the choice of specification. Specifically, the size of the estimated coefficient is particularly sensitive to the addition of variables representing physical and human capital accumulation. When these variables are added, the estimated coefficient falls by ½. The authors interpret this result as indicating that the negative effect of inflation on growth occurs partly through an investment channel where inflation reduces the rate of capital accumulation. However, this possibility is not tested for (by, for example, estimating the relationship between investment and inflation). Another problematic feature of the regression is that the time trend intended to capture productivity growth has the wrong sign.

The estimated relationship is also sensitive to the addition of fixed effects (country-specific constants) in the panel regression. When these are added, the time trend has the right sign, but the positive impact of physical and human capital accumulation on growth disappears. This suggests that the statistically significant effect of these latter two variables may be spurious in the sense that their significance in the original specification is due to the inappropriate restriction that the country specific constants are equal across countries. In the paper, the authors acknowledge this possibility but do not test for it. Finally, another results that suggests that the model may be subject to robustness problems is that the estimated relationship is sensitive to whether outliers such as Iceland and Turkey are included.

Finally, the results provide only weak support for the theoretical model, since a number of variables other than inflation that the model suggests should be included in the specification are not significant. This weak empirical support for the model limits the extent to which it can help address the problems of interpretation noted above.

An analysis of fiscal policy in the Federal Reserve Board's global model

David Bowman and John H. Rogers[1]

Introduction

The march toward European Monetary Union (EMU) following the Maastricht accord has put the issue of budget deficit reduction and government debt levels on the front-burner of policymaking in 1997. The outlook on the fiscal front presents several challenges to monetary policy. Because of the lags associated with changing monetary policy instruments, it is incumbent upon the monetary authorities to be forward-looking, especially with regard to gauging the effects of impending fiscal policy developments.

The purpose of this paper is to evaluate the effects of fiscal policy from the perspective of a large-scale, multi-country simulation model. We consider three main issues: achieving budget balance through *spending cuts versus tax increases*; permanently reducing the *target debt-to-GDP ratio*; and *imperfect credibility* of announced fiscal policies;. We also consider the effects of two alternative monetary policy rules for a particular fiscal shock. Finally, we examine a fiscal policy scenario that is motivated by impending future developments, in particular actions to keep debt and deficit levels at or below those required for participation in EMU.

Recent changes in the design of the Federal Reserve Board's global model make it particularly well-suited for the task at hand. Such changes include the explicit incorporation of intertemporal budget constraints for the government and external sectors, and the forward-looking behavior of several important variables.

All simulations reported in this paper were carried out using the joint FRB/US-FRB/MCM model. Section 1 describes the general features of the models, focusing on the FRB/MCM model which is a more integral part of our analysis. The simulations and their results are described in detail in Section 2.

1. General features of the model

1.1 FRB/MCM[2]

The FRB/MCM is a dynamic global economic model, comprised of twelve country/regional sectors with a total of nearly 1,400 equations. Each of the G-7 economies (Canada, France, Germany, Italy, Japan, the United Kingdom and the United States) is represented by about 35 behavioral equations and 100 accounting identities. The specification of these equations is fairly similar for all seven countries; the differences are mainly with respect to the estimated regression

---

[1] The authors are economists with the Division of International Finance of the Board of Governors of the Federal Reserve System, and wish to thank Andrew Levin and Ralph Tryon for useful suggestions, and Asim Husain for invaluable research assistance. The views expressed in this paper are those of the authors and should not be interpreted as representing the views of the Federal Reserve Board of Governors or other members of its staff.

[2] The material in this section is taken from Levin (1996).

coefficients and bilateral trade weights. Three sectors – Mexico, the newly industrializing economies (NIEs), and other OECD economies (ROECD) – are modelled on a more aggregated and stylized basis, with about 20 behavioral equations and 75 accounting identities each. Finally, a total of about 45 equations are used to represent the behavior of OPEC members and of other developing and transition economies (ROW).

1.1.1 Long-run stability and dynamic adjustment

The FRB/MCM is designed to exhibit long-run stability and balanced growth, similar to that of a standard neoclassical growth model. Each consumption equation incorporates error-correction mechanisms to ensure that the level of consumption (in natural logarithms) is cointegrated with disposable income and the real interest rate. Hence, conditional on the long-term real interest rate, the savings rate is stationary (perhaps around exogenous trends related to demographics or other factors). Similarly, imports are cointegrated with domestic absorption and the real exchange rate, while exports are cointegrated with foreign absorption and the real exchange rate.

The long-run stability of the FRB/MCM is also facilitated by explicitly incorporating stock-flow relationships for physical capital and present-discounted-value constraints for government and net external debt. Thus, private investment exhibits short-run accelerator-type effects in response to output fluctuations. In the longer run, however, the investment rate adjusts to equate the marginal product of capital to its real rate of return. This adjustment effectively serves as an error-correction mechanism, ensuring that the level of investment and the capital stock are each cointegrated with gross output and with the long-term real interest rate (net of depreciation).

An important feature of the model for the purposes of our paper is long-run fiscal solvency. This is maintained by an endogenous tax rate reaction function, which adjusts the income or sales tax rate when the nominal government debt/GDP ratio deviates from a specified target. In the FRB/MCM, government expenditures and tax revenues are subject to cyclical movements as well as exogenous shocks. Since budget deficit fluctuations affect the stock of debt and hence subsequent interest payments, the tax rate adjustment must be sufficiently large to prevent an explosive path of government debt. Thus, given an appropriate specification of the tax rate reaction function, the model ensures that the stock of government debt is cointegrated (in natural logarithms) with nominal GDP.

Finally, changes in the net external debt/GDP ratio lead to corresponding movements in the sovereign risk premium. Thus, through uncovered interest parity, a deterioration of the current account induces an increase in the domestic real interest rate and/or a depreciation of the real exchange rate. A reasonable degree of sovereign risk premium adjustment ensures that improved net exports of goods and non-factor services will outweigh the higher net factor payments resulting from the initial increase in external debt, and thereby prevents an explosive path for the current account and net external debt.

1.1.2 Treatment of expectations

The explicit treatment of expectations has played an important role in the formulation of the FRB/MCM. In all countries/sectors except OPEC and ROW, expected values of future variables directly influence the determination of interest rates, consumption and investment expenditures, the aggregate wage rate, and the nominal exchange rate. First, the long-term nominal interest rate and long-term expected inflation rate are each determined as geometric weighted averages of future short rates. Secondly, consumption, residential investment and business fixed investment each depend on the *ex ante* long-term real interest rate (the long-term nominal interest rate less expected inflation), while business and petroleum inventory investment each depend on the *ex ante* short-term real interest rate. Thirdly, the aggregate nominal wage rate is defined in terms of the current and past values of overlapping four-quarter wage contracts, where each wage contract depends on expected future aggregate wages and expected deviations of unemployment from its natural rate. Finally, each bilateral nominal exchange rate (local currency/US$) is determined by uncovered interest parity; i.e.

the expected rate of depreciation depends on the current bilateral interest rate differential, adjusted by the endogenously determined sovereign risk premium described above.

The FRB/MCM can be simulated under two alternative assumptions about expectations formation: VAR-based expectations (referred to as backward-looking or "adaptive" expectations), and model-consistent expectations (also referred to as forward-looking or "rational" expectations). Since assumptions about expectations formation can have important implications for the simulation results, it is useful to review the implementation of these assumptions in some detail.

The implementation of *adaptive expectations* formation in the FRB/MCM closely parallels the approach followed in the FRB/US quarterly model. In particular, regression equations have been estimated for each of the G-3 economies (Germany, Japan and the United States) using historical data on the output gap (i.e. the deviation of real GDP from potential), the GDP price deflator, the short-term Treasury bill rate and the average wage rate. The current output gap and the current price inflation rate are each regressed on up to eight quarters of lagged output gaps, inflation rates and interest rates; and the wage inflation rate is regressed on its own lags as well as lags of the other three variables.

For a given simulation experiment, a monetary policy rule must also be specified, in which the short-term interest rate is determined as a linear function of the current output gap and the rate of price inflation; e.g. the rule analysed by Taylor (1993). The interest rate reaction function is combined with the reduced-form output gap and price inflation equations to create a three-variable VAR model. For any forecasting horizon $N \geq 0$, the VAR model can be evaluated recursively to obtain a forecasting equation for each variable, in which the N-step-ahead forecast is expressed in terms of the current and lagged values of all three variables. An algorithm developed by David Bowman is used to compute the geometric weighted average of these forecasts over all horizons, yielding reduced-form equations for the long-term nominal interest rate and long-term expected inflation in terms of the current and lagged values of the output gap, the inflation rate and the short-term interest rate.

In each period of a dynamic simulation, current and lagged variables are used to evaluate each reduced-form equation and obtain new expectations of future variables. For example, the reduced-form price inflation equation is used to determine short-term expected inflation, which is needed to calculate the *ex ante* short-term real interest rate for each of the inventory investment equations. The reduced-form equations for the long-term interest rate and long-term expected inflation are used to calculate the *ex ante* long-term real interest rate, which enters the consumption, fixed investment and uncovered interest parity equations. Finally, the aggregate wage rate is determined directly from the reduced-form wage equation.

For each dynamic simulation of the FRB/MCM, *model-consistent expectations* are implemented by obtaining the perfect foresight solution path for all endogenous variables. To understand how this solution is obtained, it is useful to define the set of "expectations variables" as those endogenous variables whose expected future value enters into one or more equations in the model. The solution algorithm requires the long-run stability of all expectations variables: i.e. after a shock occurs, each expectations variable must eventually return to the baseline (or to some other known steady-state value). In this case, the baseline or steady-state values can serve as terminal conditions for the expectations variables at some date sufficiently far into the future. Thus, the perfect foresight solution algorithm determines the paths of all endogenous variables over the simulation period, using prespecified values for the terminal conditions as well as for the initial conditions and the exogenous variables.

For example, suppose that one wishes to evaluate the effects of an exogenous change in government spending over the period 1996Q1-1999Q4. If the model is reasonably stable, one might expect that all variables would return to baseline within about 25 years. Thus, the use of model-consistent expectations would typically require a dynamic simulation over the period 1996Q1-2025Q4. In this case, the required initial conditions would be the pre-1996Q1 values of all lagged variables in the model, which can be specified using historical data and/or an extrapolated baseline. The required terminal conditions would be the post-2025Q4 values of all expectations variables in the model, which would be specified based on the long-run properties of these variables.

1.1.3 Monetary policy rules

In the FRB/MCM, monetary policy is modeled using an interest rate reaction function. Many of the simulations below are run assuming that monetary policy follows the rule analysed by Taylor (1993). This rule adjusts the short-term interest rate based on deviations of inflation from its target rate and on deviations of output from potential:

Taylor's rule: $i \equiv \bar{r} + \pi^* + 1.5INFGAP + 0.5YGAP$ where *INFGAP* is defined as the deviation of current inflation from its target rate, $\pi - \pi^*$, and *YGAP* is the deviation of current GDP from potential. Taylor calculated the US federal funds rate implied by this rule and found that the implied interest rate followed a path quite similar to that of the actual federal funds rate over the period 1983-92.

If both current and expected inflation are at the target rate, and output is at potential, then Taylor's rule implies that the *ex ante* real interest rate is at the equilibrium rate, yielding steady inflation and sustainable real GDP growth. If current inflation exceeds the target rate by one percentage point, Taylor's rule prescribes a 1.5 percentage point increase in the nominal interest rate, which will typically raise the *ex ante* short-term real interest rate by about 50 basis points. (The exact increase in the *ex ante* real interest rate depends on short-term expected inflation, but this is typically quite close to the current inflation rate.) The increase in the real interest rate dampens economic activity, thereby depressing employment and placing downward pressure on wages and prices until inflation returns to its target rate.

Taylor's rule also indicates that the federal funds rate should be adjusted in response to deviations of output from potential. When economic activity is relatively weak, this component of Taylor's rule reflects the effect of an interest rate cut in stimulating economic activity. In some of the simulations that follow, we consider a variant of Taylor's rule, in which the coefficient on *YGAP* is raised to 10. This alternative prescribes a much stronger response of interest rates to a deviation of output from potential and thus, all else constant, has the effect of keeping real GDP closer to baseline.

1.2 FRB/US

The FRB/US model is described in detail in Brayton and Tinsley (1996). FRB/US is similar in spirit to an individual country sector of the FRB/MCM in terms of reliance on long-run equilibrium conditions, treatment of expectations, and modeling of dynamic adjustments. However, FRB/US, with roughly 280 equations, contains much more detail than the US model of FRB/MCM. Much of the additional detail is contained in the financial sector of the model. Bridging the FRB/US model with the foreign sectors of the FRB/MCM produces the Federal Reserve Board's global model, which is used for all simulations in this paper.

2. Simulation results

This section reports the results of four different fiscal experiments, each of which is run in the global model. We make use of both VAR-based expectations and model-consistent expectations in the various simulations. The simulations are designed to mimic several aspects of real world fiscal policy, and thereby provide some guidance for monetary policy. The simulations also highlight several new mechanisms of the FRB global model.

2.1 Government spending cuts versus tax increases

The first experiment is designed to compare the effects of a cut in government spending, arbitrarily set at 1% of GDP, to an equivalent increases in personal income taxes. So, beginning in 1997Q1, government spending is cut (or tax revenues are increased) by 1% of GDP for 7 years, after which spending (taxes) return gradually to baseline. Monetary policy is assumed to follow Taylor's rule in each country except France, whose currency is tied to the DM, and Mexico, where the currency is tied to the US dollar. The experiment is run under both VAR-based, or "adaptive", expectations and model-consistent ("rational") expectations. We are mainly interested in the magnitude of the short-run and intermediate-run effects of the shock on GDP, although we also display results for the components of GDP as well as nominal short-term and real long-term interest rates. In order to make the spending cut and tax increase scenarios as comparable as possible, we "turn off" the tax rate reaction function in the former case.

The results for Germany are depicted graphically in Figures 1a and 1b for, respectively, the adaptive and the rational expectations version. A cut in government spending directly reduces aggregate demand, and leads to a contraction of GDP that is larger on impact than a tax increase, which lowers disposable income and reduces aggregate demand indirectly through lower consumption expenditures. The monetary authority responds to the drop in real GDP by lowering short-term interest rates. The associated decline in real long rates eventually spurs investment spending and, along with the rise in net exports that is due to a drop in domestic absorption, brings GDP back to baseline. In the intermediate-run, as early as approximately six quarters after the shock, the effect on GDP is the same under either spending cuts or tax increases. Finally, a comparison of Figures 1a and 1b indicates that the results are very similar for either assumption about expectations. In both cases, (a) the peak drop in GDP is nearly twice as large for spending cuts as tax increases, and (b) the effects on GDP are the same after about six quarters.

2.2 Alternative monetary policy rules

Next we consider how the simulated effects of a particular fiscal shock may be affected by altering the monetary policy rule. We consider a permanent one percent of GDP cut in government spending as in the first simulation.[3] We focus on the rational expectations model because it best illustrates the mechanisms through which contractionary fiscal policy may have a positive effect on GDP. The potential for such an unconventional effect has recently captured a lot of attention in both academic and policy-making circles. Two alternative rules for monetary policy are considered. The first is a Taylor rule, as described above. The second rule assigns a weight of 10 to the output gap in the variant of Taylor's rule described above, while keeping the coefficient on *INFGAP* unchanged (one might loosely refer to this as a real GDP target).

The results are depicted for Germany in Figure 2. The cut in government spending has a direct negative effect on aggregate demand, and immediately moves GDP below baseline. The monetary authority responds to the output gap by lowering short rates. This response is larger under the modified Taylor rule (with a large weight on the GDP gap), as indicated by the dashed lines. Thus, as depicted at the bottom of Figure 2, nominal short rates fall by more under the real GDP rule than under the Taylor rule (almost three times as much), as do real long rates. The extra drop in long rates provides an additional boost to investment spending, props up consumption more, and hence keeps GDP closer to baseline.

---

3 In this case, however, we turn the tax rate reaction function on immediately and leave it on throughout the simulation.

Figure 1a

Government spending cuts versus tax increases (AE)

Key variables (B): Germany

Absolute (+/–) and relative (%) deviations from baseline, in percentage points

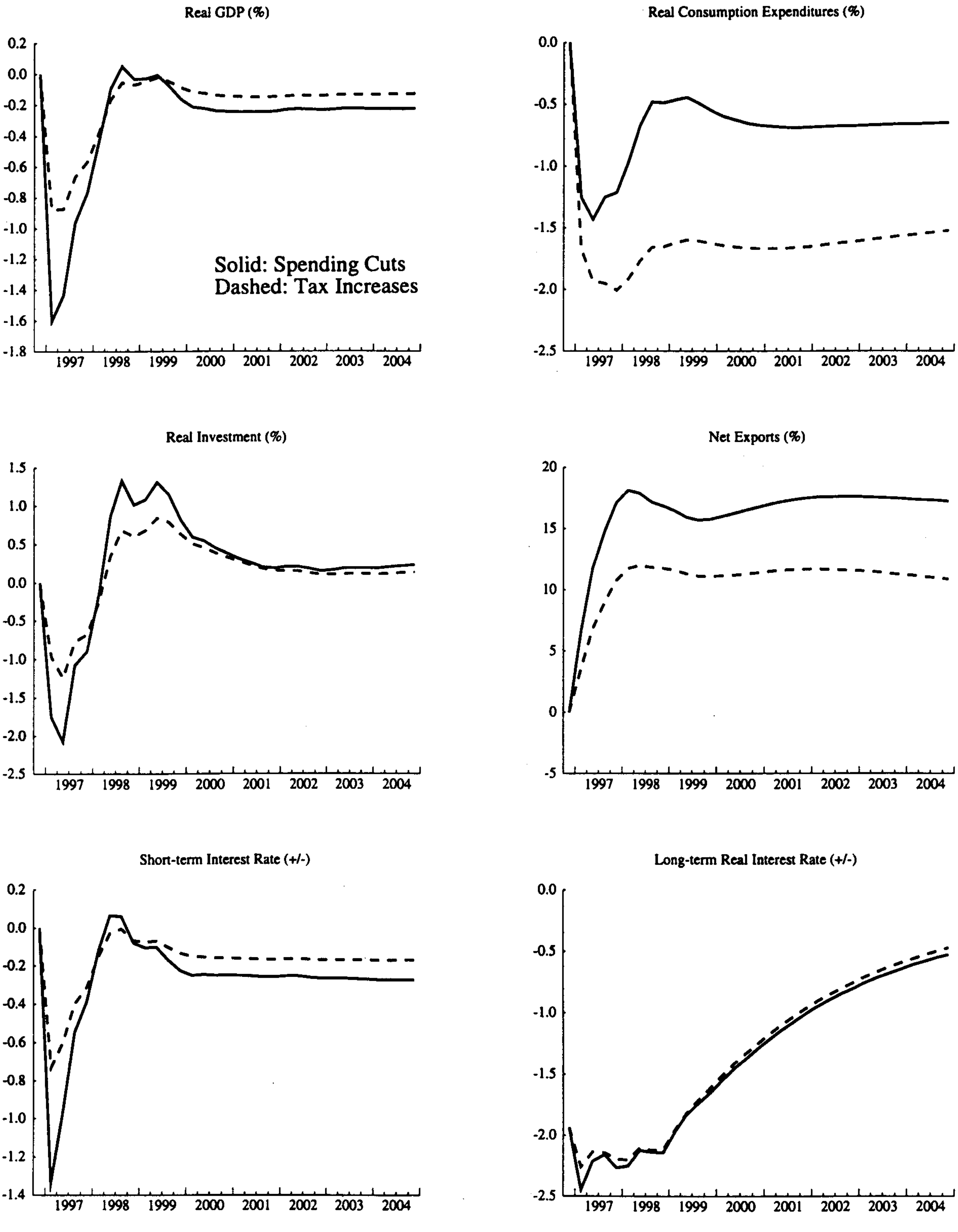


Figure 1b

Government spending cuts versus tax increases (RE)

Key variables (B): Germany

Absolute (+/–) and relative (%) deviations from baseline, in percentage points

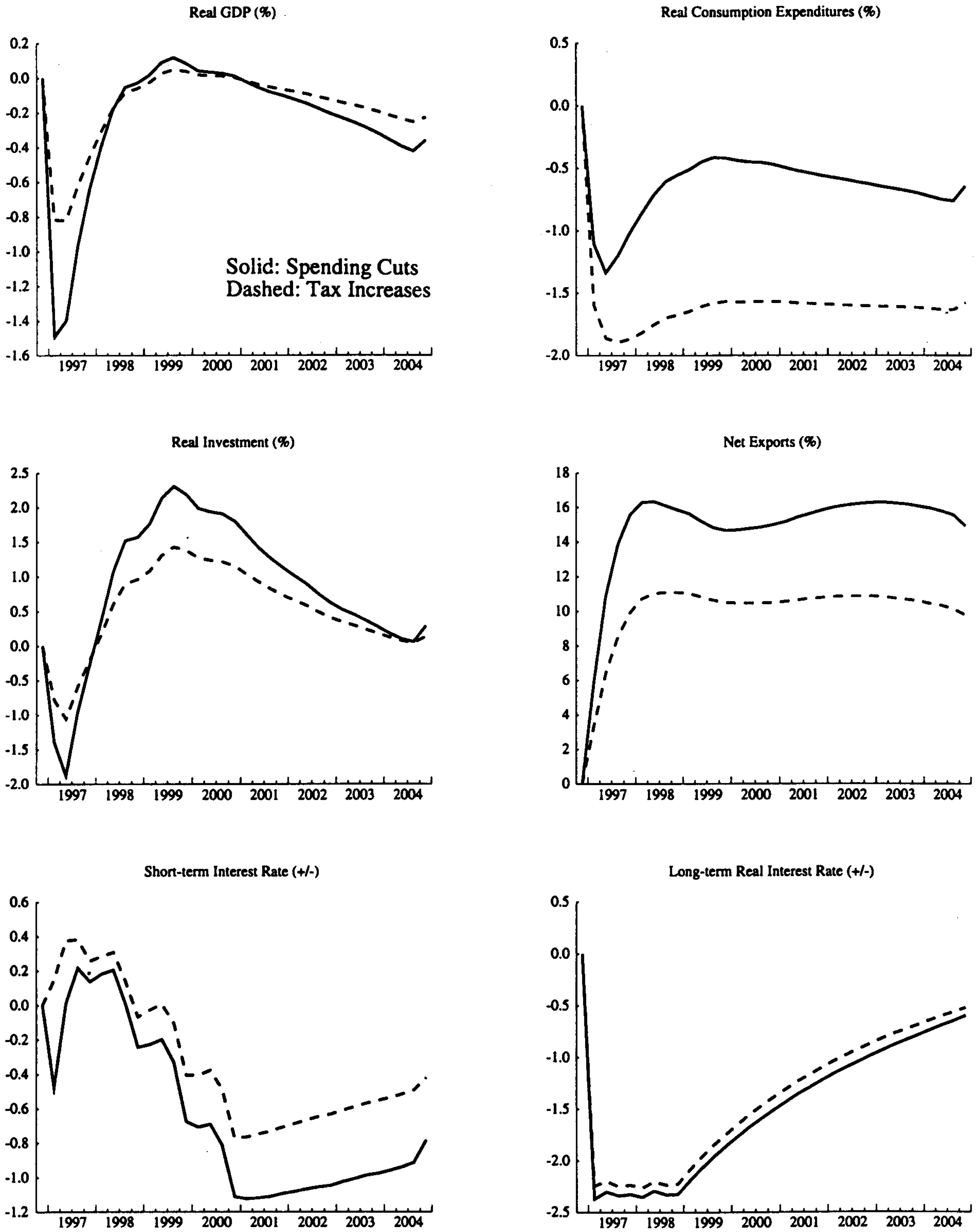


Figure 2

Comparing monetary policy rules

Key variables (B): Germany

Absolute (+/–) and relative (%) deviations from baseline, in percentage points

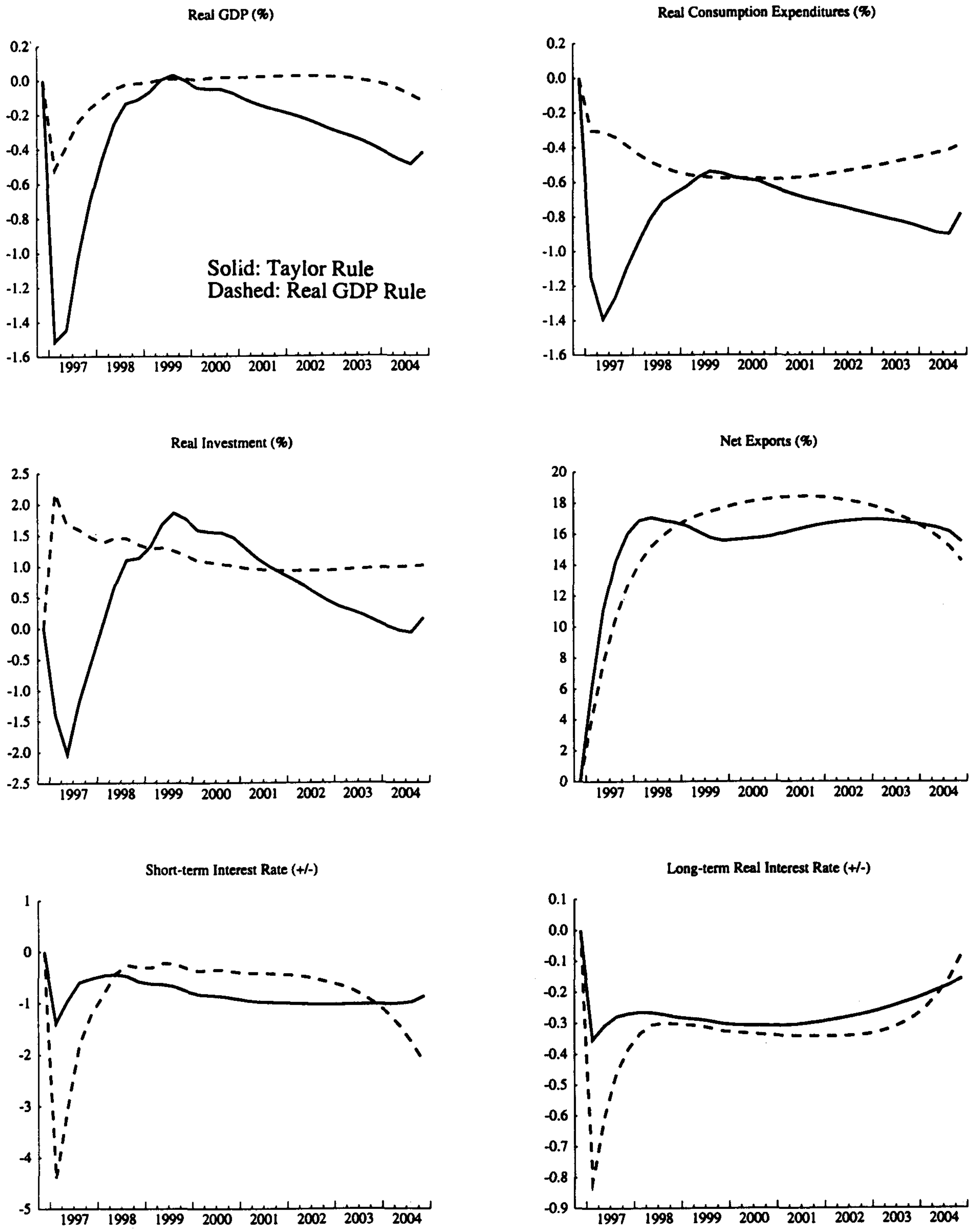
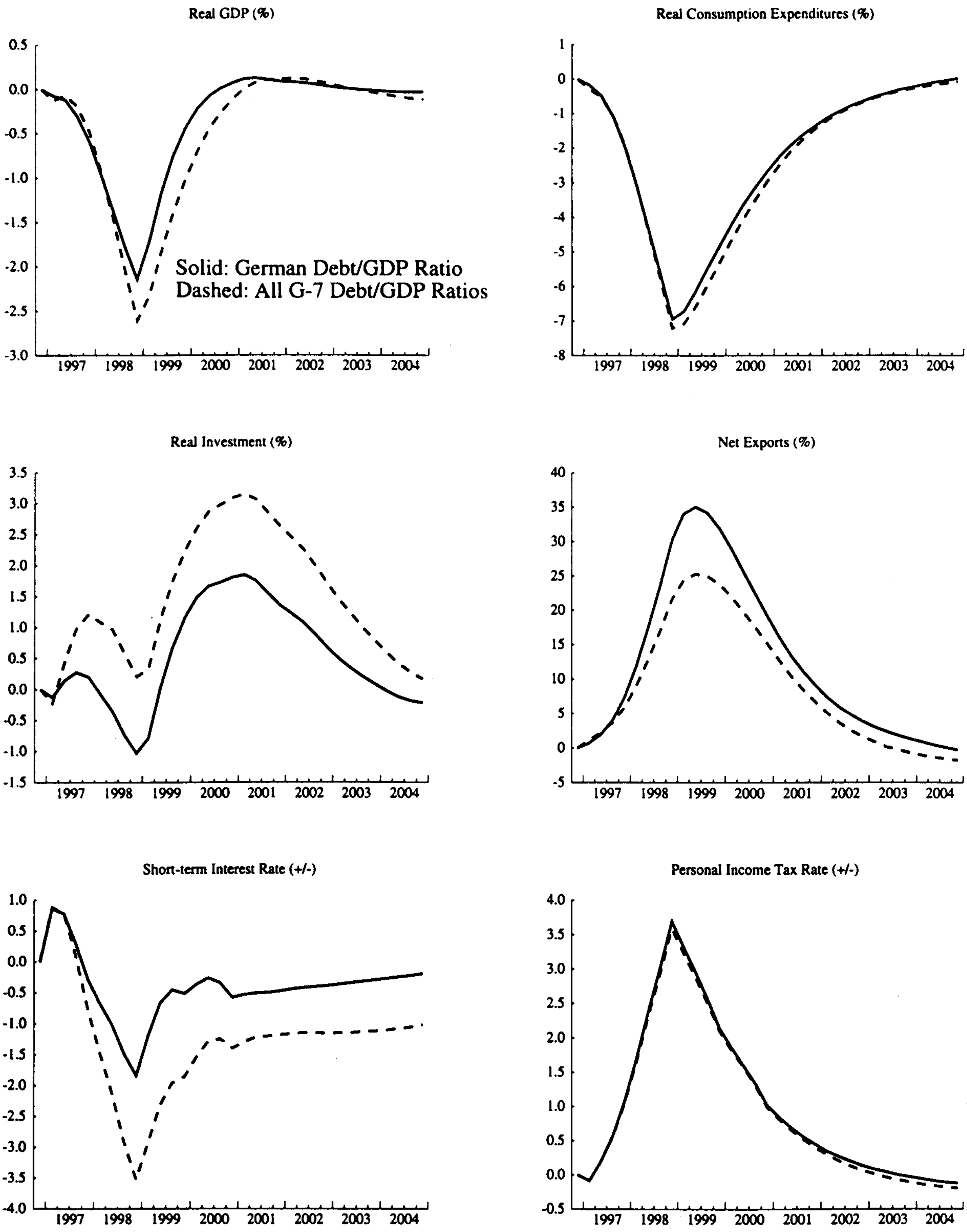


Figure 3
Reducing the target debt/GDP ratio
Key variables (A): Germany

Absolute (+/–) and relative (%) deviations from baseline, in percentage points



2.3 Permanent lowering of the target debt/GDP ratio

Recall from Section 1 that the tax rate reaction function, through which long-run fiscal solvency is maintained, adjusts the income tax rate when the debt/GDP ratio deviates from its target. In this experiment, we consider the effects of a permanent reduction of the target debt/GDP ratio to one-half its current level. Starting in 1997Q1, the target debt ratio is gradually reduced for 8 quarters, at which time it is 50 percent below its original value, where it remains permanently. We consider two cases, one in which Germany alone adjusts the target debt ratio, and a second in which each of the non-US G-7 countries adjusts the debt ratio. Only the rational-expectations version of the model is used. Other aspects of the simulation, including the assumptions about monetary policy, are the same as those in the first experiment.

Results are displayed for Germany in Figure 3. The solid line represents the case in which Germany alone reduces the debt ratio, while the dashed line represents the multi-country scenario. Notice from the bottom right panel that the shock induces a (gradual) increase in the tax rate which peaks eight quarters after the shock. The tax hike induces a drop in disposable income, and hence consumption expenditures, and a decline in GDP which also peaks at the eight-quarter point. The recession induces a loosening of monetary policy, which eventually leads to a rise in investment spending and, via a real depreciation, an improvement in net exports. Real GDP returns to baseline approximately 3 to 4 years after the shock, while the stock of government debt is permanently lower.

Finally, a comparison of the dashed and solid lines in the figure indicates that the qualitative features of the simulation described above hold for either the single-country or the multi-country case. The effects (on Germany) are slightly stronger in the case in which other G-7 countries also reduce the debt ratio, as the recessionary impact of such a policy in Germany's trading partners spills over via lower German net exports.

2.4 EMU participation scenario

Finally, we examine a fiscal policy scenario that is motivated by likely future developments, in particular actions to keep debt and deficit levels at or below those required for participation in EMU. We also use this simulation to gauge the extent to which past and prospective monetary policy actions are likely to be sufficient to offset the negative effects associated with the fiscal actions.

A detailed breakdown of the fiscal shock is provided in the following table. Each fiscal shock was implemented at an equal amount each quarter within the year. The 1997 (1998) changes were added cumulatively on top of the cuts which occurred in the previous year(s). After 1998, all fiscal variables gradually return to baseline by 2006. The tax reaction function for each country was shut off until 2001Q1. The assumed monetary policy rule is that which puts a weight of 10 on the output gap in the interest rate reaction function. Once again, Mexico ties monetary policy to the United States and France ties policy to Germany.

Simulations were run using both expectational assumptions. Because we also wish to use this scenario to examine the effects of imperfect credibility, we focus only on the model-consistent expectations version, as this model contains forward-looking elements. The experiment is conducted in two parts. In the first, policy is completely credible: the entire path of fiscal policy changes is correctly foreseen. In the second, policy is imperfectly credible, in the sense that the 1998 policy changes are not anticipated as of 1997.

Simulations were run using both expectational assumptions. Because we also wish to use this scenario to examine the effects of imperfect credibility, we focus only on the model-consistent expectations version, as this model contains forward-looking elements. The experiment is conducted in two parts. In the first, policy is completely credible: the entire path of fiscal policy changes is correctly foreseen. In the second, policy is imperfectly credible, in the sense that the 1998 policy changes are not anticipated as of 1997.

Summary of fiscal shock used in EMU-inspired simulation

	Fiscal changes in 1996 (as a percentage of GDP)		
	Taxes	Goods & services	Transfers
Canada	0.117	-0.927	-0.126
France	0.348	-0.568	-0.244
Germany*	-0.050	0.050	0
Italy*	-0.485	0.121	0.364
Japan*	0	0.360	0.040
United Kingdom	0	-1.350	0

	Fiscal changes in 1997 (as a percentage of GDP)		
	Taxes	Goods & services	Transfers
Canada	0.028	-0.222	-0.030
France	0	-0.371	-0.159
Germany	0.212	-0.424	-0.424
Italy	0.335	-0.084	-0.251
Japan	0.720	-1.404	0.324
United Kingdom	0	-0.650	0

	Fiscal changes in 1998 (as a percentage of GDP)		
	Taxes	Goods & services	Transfers
Canada	0.039	-0.351	0
France	0	-0.315	-0.135
Germany	0	-0.090	-0.090
Italy	0.420	-0.105	-0.315
Japan	0.160	-0.312	0.072
United Kingdom	0	-0.560	0

* Denotes an expansionary policy.

Figure 4 depicts the results graphically for Germany. The hike in taxes and cut in government spending, which total about 1% of GDP in 1997, have a contractionary effect on output.[4] Under the interest rate rule used, the monetary authority responds by cutting short-term interest rates. In the case of perfect credibility (the dashed lines), short rates are cut by less than under imperfect credibility.

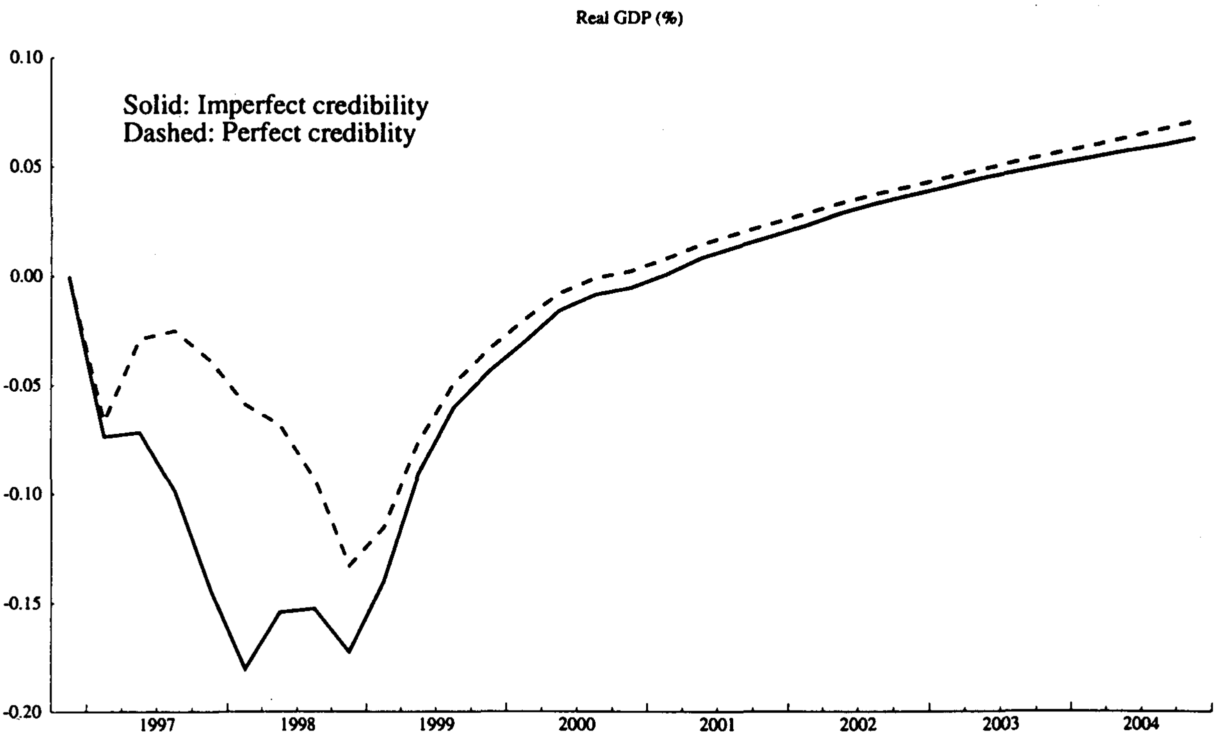
It is useful to take from this simulation some guidance concerning the appropriate degree of monetary ease that will be required to keep GDP at (or "close to") baseline, given the impending fiscal cuts.

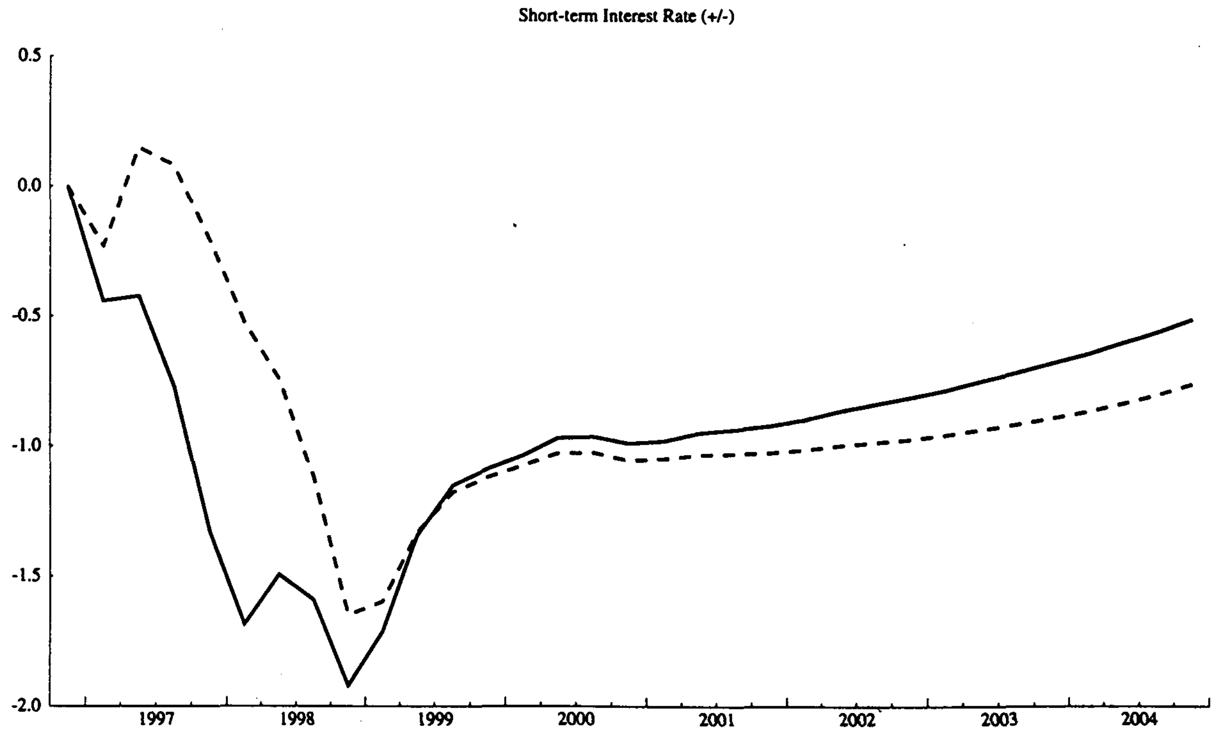
The following table displays figures for the change in real long-term interest rates over the period 1995Q4-1997 (where 1997 is the annual average rate). The first column displays the change in real long rates implied by the model simulation. The second column displays the average over this same period of the Board staff's projection of the actual change in real long rates (actual numbers for 1996 and forecasted numbers for 1997). The third column displays the difference. Based

4 The peak drop in GDP is substantially less than what was observed with either tax increases or spending cuts in Figure 1b, where a fiscal contraction equal to 1% of GDP in 1997 was also examined. There are two main reasons for this. First, in the present case there is a relatively high weight on the GDP gap in the interest rate reaction function. Secondly, in the outer years – 1998 and beyond – the fiscal contraction is smaller in this case.

Figure 4
European Monetary Union scenario
Key variables (B): Germany

Absolute (+/–) and relative (%) deviations from baseline, in percentage points





on these results we would characterize the model as saying that monetary policy will do enough to offset fiscal cuts in Canada, France, and Italy, while for the United Kingdom, factors other than monetary policy may be at work to offset the fiscal cut.[5]

Real long-term interest rates in the EMU scenario

	Change in real long rates from model (1995Q4-1997 average)	Projected change in real long rate (1995Q4-1997 average)	Difference (projected less model simulation)
United Kingdom..............	-0.89	-0.07	+0.82
Canada............................	-0.34	-0.51	-0.17
France.............................	-0.57	-1.11	-0.54
Germany..........................	-0.62	-0.18	+0.44
Italy.................................	-0.35	-3.97	-3.62
Japan...............................	-0.70	-0.18	+0.52

Conclusions

The lags associated with monetary policy make it imperative to gauge the effects of impending fiscal policies. This paper evaluates the effects of fiscal policy from the perspective of the Federal Reserve Board's global simulation model.

Four main results are worth noting. First, achieving fiscal consolidation through spending cuts has a larger short-term recessionary effects than via tax increases. Secondly, permanently reducing the debt-to-GDP ratio has a recessionary effect whose length depends critically on both the level of debt-reduction and the speed with which any particular level of debt reduction is implemented. Thirdly, the monetary policy response that is required to offset the recessionary effects of fiscal consolidation must be larger the less credible is the announced fiscal package. Finally, using a fiscal policy scenario motivated by impending actions to keep debt and deficit levels at or below those required for participation in EMU, we project that monetary policy will do enough to completely offset the recessionary effects of fiscal cuts in Canada, France, and Italy.

References

Brayton, F. and P. Tinsley (ed.), (1996): "A Guide to FRB/US". FEDS *Working Paper* No. 96-42, Board of Governors of the Federal Reserve System.

Levin, A. (1996): "A Comparison of Alternative Monetary Policy Rules in the FRB Multi-Country Model", Board of Governors of the Federal Reserve System.

Taylor, J.B. (1993): "Discretion vs. policy rules in practice". *Carnegie-Rochester Conference Series on Public Policy*, No. 39, pp. 195-214.

---

[5] Japan is a counter-intuitive case, as the model appears to be predicting negative nominal interest rates in light of the large fiscal contraction in 1997.

Comments on: "An analysis of fiscal policy in the Federal Reserve Board's global model" by David Bowman and John H. Rogers

by Frank Smets

The question of the recessionary impact of fiscal consolidation and the appropriate monetary policy response is an important and topical one, in particular in the context of the drive towards EMU and the need to fulfil the fiscal convergence criteria of the Maastricht Treaty. This paper uses the Federal Reserve Board's Global Model to simulate various fiscal policy scenarios. The main results are that, first, spending cuts are more recessionary than tax increases and, second, that the recessionary impact of permanently reducing the level of debt will depend on the level of debt-reduction and the monetary policy response. In particular, the authors find that the recessionary effects of the additional fiscal tightening necessary to fulfil the Maastricht criteria will be limited if monetary policy eases sufficiently.

These results are familiar in that they can be derived in a standard IS/LM model and are similar to the simulation results obtained by the OECD using the Interlink model and by the IMF using the Multimod model. They contrast, however, with the analysis of episodes of fiscal consolidation by, for instance, Giavazzi and Pagano, Alesina and Perotti and Wescott and McDermott. These authors find that large packages based on spending cuts may be less recessionary than small packages which rely on tax increases. To explain these results one needs to rely on the positive effects of fiscal consolidation programmes on consumption and investment through higher expectations of future output and lower risk premia.

Although the Global Model of the FRB does include forward-looking expectations and a country risk premium, the results are more in line with the standard IS/LM analysis than with some of the historical episodes of fiscal consolidation. It would be interesting if the authors included a discussion of the various channels that may be operative and which ones are captured in the FRB model. One feature of the simulations which may limit the importance of the expectational channels is the specification of the steady state. First, in most simulations it has a fixed and known debt/GDP steady-state ratio and a fixed and known inflation target. Many of the risk premia and expectational effects of fiscal consolidation have to do with a reduced fear of monetisation and/or the effects on the credibility of the central bank's inflation target. Because of the fixed and credible inflation and debt targets, such effects are absent from the model. Second, one would like to explore to what extent different consolidation programmes (in terms of taxes versus spending cuts, government consumption versus investment cuts) have different effects on long-run output and to what extent these long-run effects can feed back into current consumption and investment. Finally, one element that is operative in the FRB model is the effect of debt on the risk premium. There is, however, no indication in the paper how important this effect is in offsetting the short-run negative demand effects of fiscal consolidation.

A more significant role for credibility effects may not only reverse some of the results concerning the output effects of fiscal consolidation, it may also affect how monetary policy should respond to fiscal consolidation. For example, the paper finds that because a less credible package has a larger recessionary impact, the central bank needs to ease policy more the less credible the package. Central bankers often have a different view: limited credibility of the fiscal package requires a cautious response or the credibility of the central bank may be negatively affected.

One of the comparative advantages of using a multi-country model is that one can explicitly address the role of the exchange rate. It is, therefore, somewhat surprising that the authors do not address this issue. The role of the exchange rate may be important in a number of ways. First, when comparing single-country with multiple country fiscal consolidations, the usual crowding-in through exchange rate depreciation may be less. Is this the case in the FRB model and what are the implications for the monetary policy response? Second, the effects on the exchange rate are important in assessing what happens to risk premia as a result of fiscal consolidation. Finally, the monetary policy reaction function is only in terms of output and inflation. In open economies the monetary authorities may also respond to the exchange rate.