Endogenous Monetary Policy with Unobserved Potential Output*

Alex Cukierman† and Francesco Lippi‡

February 28, 2005

Abstract

This paper characterizes monetary policy when policymakers are uncertain about the extent to which fluctuations in output and inflation are due to changes in potential output or to cyclical demand and cost shocks. Our results suggest an explanation for the inflation of the seventies and the price stability of the nineties. It is shown that: (1) policy is likely to be excessively loose for some time when there is a large decrease in potential output in comparison to a full information benchmark. (2) Retrospective policy errors and errors in forecasting potential output and the output gap are generally serially correlated. (3) The increase in the Fed's conservativeness between the seventies and the nineties implies that the information problem had greater consequences in the former period.

JEL Classification Numbers: E5

Key Words: monetary policy, potential output, filtering, inflation, output gap.

---

*Earlier versions of the paper were presented at the November/December 2001 NBER conference on "Macroeconomic policy in a dynamic, uncertain economy", at the Bank of Italy, the Board of Governors of the Fed, the Bank for International Settlements, the Central European University, the European Central Bank, the Free University of Brussels, the University of Cologne and the UCSC. We benefitted from extremely useful discussions with Athanasios Orphanides, from the comments of Simon Gilchrist, David Small and Philippe Weil and from the suggestions of two anonymous referees and the editor.

†Berglas School of Economics, Tel-Aviv University and CEPR. e-mail: alexcuk@post.tau.ac.il.

‡Research Department, Bank of Italy.

1. Introduction

A stabilizing role for monetary policy hinges on some notion of "potential output", a non-observable economic variable that is central for the determination of the target level of output. The conduct of monetary policy therefore requires that the central bank estimates and continually updates, its measure of potential output. Kuttner (1992, 1994) was among the first to raise the issue of the quantitative importance of uncertainty about potential output for real-time policymaking. He examined the difficulties inherent in real-time estimation of potential output and suggested that situations requiring policy actions might not be immediately recognizable because of signal extraction errors arising under imperfect information.

This policy implication is central for Orphanides (2001,2003a,b), who reports evidence of a significant (real time) overestimation of potential output during the oil shocks of the seventies. Enlightening documentation on the ex-post downward revisions of potential output appears in the 1979 Economic Report of the President (Chart 7, pp. 72-76), reported below, which vividly illustrates the magnitude and persistence of the revisions.

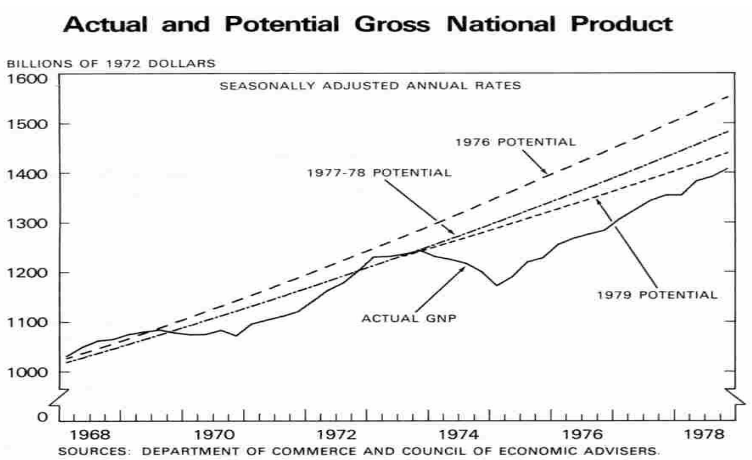

Figure 1 - Reprint from "Economic Report of the President (1979)".

Orphanides argues that by leading to a monetary policy stance which turned out to be, with the benefit of hindsight, excessively loose, the real-time overestimation of potential output aggravated inflation at the time. Somewhat symmetrically, the strong productivity gains recorded in the United States during the second half of the 1990s raise the possibility that the greater-than-expected increases in potential output could have allowed a less restrictive monetary policy stance than the one implied initially by real time estimates of the output gap and inflation.

The hypothesized relevance of imperfect information may shed new light on monetary policy "errors" during the seventies and raises an important question about the extent to which such retrospective policy mistakes can be avoided in the future. If the errors were due to poor forecasting procedures or to an inefficient specification of the "policy rule", a likely answer to this question is yes. But if, given the available real-time information, policy was as efficient as possible, the likely answer is no. Assessing the extent to which retrospective policy mistakes are due to "bad policies" rather than to "bad luck" requires a model which identifies optimal monetary policy under imperfect information. The availability of such a benchmark is essential for evaluating the extent to which (retrospective) policy mistakes were avoidable in real time. This paper makes a step in this direction by analyzing such a benchmark model.

We embed the real-time information problem into a simple backward looking macroeconomic model by assuming that the central bank cannot perfectly distinguish (not even ex-post) between fluctuations in inflation and output that are due to shocks in potential output and those that are due to higher frequency demand and cost shocks. We label this inevitable confusion as the "information problem", IP in brief.[1] To isolate the effects of uncertainty that arise from the IP, we adopt a specification that features the certainty equivalence property of optimal policy so that the form of the policy function in terms of optimal forecasts of relevant variables is invariant to uncertainty. But the mapping from real time information into those forecasts and, therefore, macroeconomic outcomes, do depend on uncertainty. The main purpose of the paper is to study how the IP influences the dynamics of these outcomes, in particular output and inflation, in comparison to a full information benchmark.

The results show that, given the structure of information, some policy decisions that are judged ex-post to be "mistakes" may be unavoidable in real-time even

[1]The macroeconomic consequences of a similar confusion were discussed following the oil shocks of the seventies within frameworks in which monetary policy is exogenous (Brunner, Cukierman and Meltzer (1980), Part II of Cukierman (1984) and Chapter 4 of Brunner and Meltzer (1993)). This earlier literature referred to the inability to perfectly distinguish between permanent and transitory shocks to productivity as a "permanent - transitory" confusion.

if the central bank uses the best forecasting procedures. These retrospective mistakes are normally small during periods in which changes in potential output are small. But during periods characterized by unusually large changes in the long run level of output, policy mistakes in a given direction are likely to be large and to persist for some time.[2]

The evidence in Orphanides (2001) supports the view that monetary policy during the seventies was excessively loose since a reduction in potential output was interpreted for some time as a negative output gap. This paper provides analytical foundations for this mechanism within a stylized backward looking model and identifies conditions under which the IP leads monetary policy to be *systematically* looser than under perfect information in periods of large reductions in potential output, and to be overly restrictive relative to this benchmark in periods of large expansions in potential output. The intuitive reason is that, even when they filter available information in an optimal manner, policymakers as well as the public at large detect changes in potential output only *gradually.* When there is a large decrease in potential output, as was the case in the seventies, policymakers interpret part of this reduction as a negative output gap due to insufficient demand and loosen monetary policy too much in comparison to a benchmark without the IP. Thus, in periods of large decreases in potential output, inflation accelerates partly because of the relatively expansionary monetary policy stance. Conversely, when - as may have happened in the US during the nineties - a “new economy” raises the level of potential output, inflation subsides partly because policymakers interpret some of the increase in output as a positive output gap, so that policy is tighter than under perfect information.[3]

[2]Related work in which potential output is specified as a Hodrick-Prescott filter appears in Lansing (2000). Lansing presents a more elaborate lag structure than we do, but does not derive the forecast of potential output from the stochastic structure of the economy and the policy rule from the loss function of policymakers as we do here. There is, thus, a tradeoff between his approach and ours. A paper by Swanson (2004) is nearer to our framework in that it features optimizing policymakers and specifies the estimation of potential output as a signal extraction problem. But his main point is that, in spite of quadratic objectives, the optimal policy rule depends on the variances of shocks via the solution to the signal extraction problem, whereas we focus on the implications of such a framework for optimal interest rate policy and for the associated retrospective “policy errors”.

[3]Although this mechanism may appear obvious at first blush it is by no means evident since, as demonstrated in the paper, it applies only when the persistence of demand shocks is sufficiently large in comparison to the persistence of cost shocks. In the opposite case, an

The paper shows that, even when the real-time information is processed efficiently and monetary policy chosen optimally, the forecast errors in real time estimates of potential output and of the output gap are normally serially correlated even in the population. In general, this serial correlation is induced by shocks to potential output, as well as to the cyclical components of output. The paper identifies conditions under which the bulk of the serial correlation is due to shocks to potential output. In particular, it shows that, when the variance of shocks to potential output is relatively small, most of the measured serial correlation is due to innovations to potential output. Interestingly, retrospective evidence about forecast errors in potential output during the seventies and the eighties is consistent with these implications (Orphanides, 2003b). As a consequence of the serial correlation in those errors monetary policy also appears in retrospect to be systematically biased in one direction.

In summary the paper provides a simple unified, backward looking, framework for understanding some of the reasons for both the inflation of the seventies and the remarkable price stability of the nineties and shows, by means of simulations, that similar mechanisms operate in the presence of more elaborate lag structures. It illustrates how the speed of learning by policymakers and the deviations of policy from an ideal full-information-benchmark depend on the stochastic structure of various economic shocks. Identification of such conditions is a necessary first step for gauging empirically whether imperfect information is quantitatively important. Finally, the paper argues that the IP problem is likely to have been less important during the nineties than during the seventies for two reasons. In the nineties, the Fed was more conservative and the evaluation of uncertainties surrounding potential output was more realistic.

The paper is organized as follows. Section 2 presents a simple model of endogenous monetary policy in the presence of imperfect information about the origins of fluctuations in output and characterizes optimal monetary policy in this environment. The consequences for the behavior of real interest rates, inflation and the output gap in comparison to their full information counterparts are analyzed in Section 3. Section 4 develops the real-time optimal forecast of potential output and shows that forecast errors of real time estimates of potential output and of the output gap are serially correlated. Section 5 discusses reasons supporting the

underperceived increase in potential output actually leads policy to be overly expansionary.

view that retrospective policy errors were smaller during the nineties than during the seventies. This is followed by concluding remarks. Extensive analytical derivations and simulations are relegated to the appendix.

2. Endogenous monetary policy in the presence of uncertainty about potential output

This section presents a simplified version of a backward looking sticky-price model in the presence of imperfect information about potential output. Its main advantage is that it provides a conceptually precise analytical illustration of some basic consequences of the interaction between policy and the economy in the presence of imperfect information. Using a more elaborate lag structure, the last part of the appendix shows by means of simulations that many of the qualitative conclusions obtained with this simple structure are robust to the introduction of more extended lags of the type presented in Svensson (1997).

2.1. The economy

In this framework (the logarithm of) output (y_t) and inflation (π_t) are determined, respectively, as follows:

$$y_t = z_t - \varphi r_t + g_t \tag{2.1}$$

$$\pi_t = \delta \pi_{t-1} + \lambda (y_t - z_t) + u_t, \; 1 \geq \delta > 0. \tag{2.2}$$

Here z_t denotes (the log of) potential output as of period t, r_t is a *real* short term interest rate, g_t is a demand shock and u_t a cost-push shock. The first equation states that potential output z is a fundamental long run determinant of actual output. But actual output is also affected by a demand shock and by the real rate of interest, which for given inflationary expectations, is determined in turn by the (nominal) interest rate policy of the central bank. The second equation states that, given last period's inflation, current inflation depends positively on the output gap, $y_t - z_t$ and on the cost-push shock. The lagged inflation term captures inflation persistence.

We assume the economy is subject to two types of temporary but persistent shocks and to a permanent shock to the level of potential output. The temporary shocks are the aggregate demand shock, g_t, and the cost-push shock, u_t. In line with conventional macroeconomic wisdom we postulate that the demand and cost shocks are less persistent than changes in potential output which are affected by long run factors such as technology and physical and human capital formation.[4] The persistence of shocks to potential output is modeled by assuming that z_t is a random walk.[5] More specifically we postulate the following stochastic processes for the shocks:

$$g_t = \mu g_{t-1} + \hat{g}_t \qquad 0 < \mu < 1 \tag{2.3}$$

$$u_t = \rho u_{t-1} + \hat{u}_t \qquad 0 < \rho < 1 \tag{2.4}$$

$$z_t = z_{t-1} + \hat{z}_t \tag{2.5}$$

where the innovations $\hat{g}_t, \hat{u}_t$ and $\hat{z}_t$ are uncorrelated, have zero means and respective standard deviations σ_g, σ_u and σ_z.

2.2. Monetary Policy

Policy is described by the choice of the short term real rate, r_t, made possible by the assumption of e.g. sticky prices. The policy goal is to minimize the objective function:

$$L_t \equiv \frac{1}{2} E \left\{ \sum_{j=0}^{\infty} \beta^j \left[\alpha (x_{t+j})^2 + (\pi_{t+j})^2 \right] \mid J_{t-1} \right\} \qquad \alpha > 0 \tag{2.6}$$

where $x_t \equiv y_t - z_t$ denotes the output gap (defined as the difference between the logs of actual and of potential output) and J_{t-1} is the information set available at the beginning of period t, when r_t is chosen. Substituting equation (2.1) into (2.2), the inflation equation can be rewritten as

$$\pi_t = \delta \pi_{t-1} - \lambda \varphi r_t + \lambda g_t + u_t \tag{2.7}$$

[4]The notion that demand shocks are relatively less persistent than shocks to potential output underlies the empirical identification of demand and supply factors in Blanchard and Quah (1989).

[5]Nothing in our results would change if we had added a deterministic trend growth to the potential output process.

Lagging this equation by one period to expresss π_{t-1} in terms of π_{t-2}, substituting the reulting expression into (2.7), continuing like that ad infinitum, and noting that $\delta^s \pi_{t-s}$ tends to zero as $s \to \infty$, we obtain

$$\pi_t = \sum_{s=0}^{\infty} \delta^s \left[-\lambda\varphi r_{t-s} + \lambda g_{t-s} + u_{t-s} \right]. \tag{2.8}$$

Substituting equations (2.1) and (2.8) into (2.6) to express central bank objectives in terms of the interest rate instrument

$$L_t \equiv \frac{1}{2} E \left\{ \sum_{j=0}^{\infty} \beta^j \left[\begin{array}{c} \alpha \left\{ -\varphi r_{t+j} + g_{t+j} \right\}^2 \\ + \left\{ \sum_{s=0}^{\infty} \delta^s \left[-\lambda\varphi r_{t+j-s} + \lambda g_{t+j-s} + u_{t+j-s} \right] \right\}^2 \end{array} \right] \mid J_{t-1} \right\}. \tag{2.9}$$

The first order condition for this problem with respect to the current interest rate, r_t, implies

$$-\alpha\varphi \left(-\varphi r_t + g_{t|t-1} \right) - \lambda\varphi \left[-\lambda\varphi r_t + \lambda g_{t|t-1} + u_{t|t-1} \right] \left\{ 1 + \beta\delta + (\beta\delta)^2 + ... \right\} = 0. \tag{2.10}$$

Here $g_{t|t-1}$ and $u_{t|t-1}$ are the expected values of the demand and cost shocks conditional on the information available at the beginning of period t: J_{t-1}. At this stage we note that J_{t-1} contains, among other data, observations on actual inflation and output up to and including period $t-1$. A full specification of J_{t-1} appears below. Since period t values of those shocks are not known with certainty at the beginning of period t, these variables appear in equation (2.10) in expected terms.

Using (2.10) to solve for r_t and rearranging yields, after some algebra, the interest rate rule in equation (2.11) below. The equilibrium outcome for output appears in equation (2.12) and is obtained by substituting the interest rate rule into equation (2.1). Finally, the equilibrium outcome for infation appears in equation (2.13) and is obtained by substituting (2.12) into equation (2.2)

$$r_t = \frac{1}{\varphi}\left[g_{t|t-1} + \frac{\lambda}{\lambda^2+(1-\beta\delta)\alpha}u_{t|t-1}\right] \tag{2.11}$$

$$y_t = z_t + g_t - g_{t|t-1} - \frac{\lambda}{\lambda^2+(1-\beta\delta)\alpha}u_{t|t-1} \tag{2.12}$$

$$\pi_t = \delta\pi_{t-1} + \lambda\left(g_t - g_{t|t-1}\right) - \frac{\lambda^2}{\lambda^2+(1-\beta\delta)\alpha}u_{t|t-1} + u_t. \tag{2.13}$$

2.3. The structure of information and optimal policy

The interest rate rule in (2.11) led by one period implies that the optimal real interest rate policy for period $t+1$, r_{t+1}, requires the policymaker to form expectations about the values of the demand shock and the cost push shocks, g_{t+1} and u_{t+1}. Although he does not observe those shocks directly, the policymaker possesses information about economic variables from which noisy, but optimal, forecasts of the shocks can be derived. In particular, we assume that policymakers know the true structure of the economy: $\Omega \equiv \left\{\varphi, \lambda, \rho, \mu, \sigma_u^2, \sigma_g^2, \sigma_z^2\right\}$ but do not know the precise stochastic sources of fluctuations in output and inflation.

Thus, when the interest rate r_{t+1} is chosen at the beginning of period $t+1$, the policymaker forms expectations about g_{t+1} and u_{t+1} using historical data. The latter consists of observations on output and inflation up to and including period t. The information available at the beginning of period $t+1$ is summarized by the information set

$$J_t = \{\Omega, y_{t-i}, \pi_{t-i}, \mid i = 0, 1, 2, ...\} \tag{2.14}$$

which is used to form the conditional expectations: $g_{t+1|t}$ and $u_{t+1|t}$. Past observations on output and inflation are equivalent to past observations on the two signals, $s_{1,t}$ and $s_{2,t}$ (obtained by rearranging (2.12) and (2.13)):

$$s_{1,t} \equiv y_t + g_{t|t-1} + \frac{\lambda}{\lambda^2+(1-\beta\delta)\alpha}u_{t|t-1} = z_t + g_t \tag{2.15}$$

$$s_{2,t} \equiv \pi_t - \delta\pi_{t-1} + \lambda g_{t|t-1} + \frac{\lambda^2}{\lambda^2+(1-\beta\delta)\alpha}u_{t|t-1} = \lambda g_t + u_t \tag{2.16}$$

where variables to the left of the equality sign are observed separately while those to the right are not.[6] Clearly, $s_{1,t}$ and $s_{2,t}$ contain (noisy) information on g_t and u_t which can be used to make inference on g_{t+1} and u_{t+1}, using the fact that $g_{t+1|t} = \mu g_{t|t}$ and $u_{t+1|t} = \rho u_{t|t}$.

The optimal estimates of g_t and u_t conditional on J_t ($g_{t|t}$ and $u_{t|t}$) follow immediately from the two signals (2.15) and (2.16), once the optimal estimate of potential output, $z_{t|t}$, is known.[7] Therefore, the signal extraction (or filtering) problem solved by the policymaker reduces to an inference problem concerning the level of potential output.

2.4. Mismeasurement of potential output and policymakers' views about the state of the economy

Let policy makers' forecast errors concerning the variables z_t, g_t, u_t conditional on the information set J_t be:

$$\tilde{u}_{t|t} \equiv u_t - u_{t|t} \tag{2.17}$$
$$\tilde{g}_{t|t} \equiv g_t - g_{t|t} \tag{2.18}$$
$$\tilde{z}_{t|t} \equiv z_t - z_{t|t} \tag{2.19}$$

Using (2.15) and (2.16) the following useful relationship between these errors can be derived :

$$\lambda \tilde{z}_{t|t} = -\lambda \tilde{g}_{t|t} = \tilde{u}_{t|t}. \tag{2.20}$$

The last equation shows that overestimation of potential output ($\tilde{z}_{t|t} < 0$) simultaneously *implies* an overestimation of the cost-push shock and an underestimation of the demand shock.[8] This is summarized in the following remark.

---

[6]In particular, the construction of the signals, s_{1t} and s_{2t} needed for the formation of the forecasts $u_{t+1|t}$, $g_{t+1|t}$ and $z_{t+1|t}$ utilizes the previous period forecasts $u_{t|t-1}$ and $g_{t|t-1}$, which are known at the beginning of period $t+1$.

[7]This follows from the fact that: $g_{t|t} = s_{1,t} - z_{t|t}$ and $u_{t|t} = s_{2,t} - \lambda(s_{1,t} - z_{t|t})$.

[8]This can be seen immediately by rewriting the expressions for the estimates of g and u as

$$g_{t|t} = g_t - \tilde{z}_{t|t} \tag{2.21}$$
$$u_{t|t} = u_t + \lambda \tilde{z}_{t|t}. \tag{2.22}$$

Remark 1. *Potential output overestimation* ($\tilde{z}_{t|t} \equiv z_t - z_{t|t} < 0$) *implies:*

(i) demand shock underestimation ($\tilde{g}_{t|t} \equiv g_t - g_{t|t} > 0$)

(ii) cost-push shock overestimation ($\tilde{u}_{t|t} \equiv u_t - u_{t|t} < 0$)

Inequalities with opposite signs hold when $\tilde{z}_{t|t} > 0$.

The intuition underlying this result can be understood by referring to equations (2.15) and (2.16). The first equation implies that an increase in $s_{1,t}$ is always and optimally interpreted as being due partly to an increase in z_t and partly to an increase in g_t. Similarly, an increase in $s_{2,t}$ is interpreted as being partly due to an increase in g_t and partly to an increase in u_t. Thus, when only z_t increases, part of this increase is interpreted as an increase in potential output, but the remainder is interpreted as an increase in g_t. As a consequence the error in forecasting z_t is positive and the error in forecasting g_t is negative, producing a negative correlation between the forecast errors in those two variables. Since $s_{2,t}$ does not change the (erroneously) perceived increase in g_t is interpreted as a decrease in u_t, producing a positive forecast error for this variable, and therefore, a positive correlation between the forecast errors in u_t and in z_t.

3. Consequences of forecast errors in potential output for monetary policy, inflation and the output gap

Remark 1 shows how mismeasurement of potential output distorts policymakers' perceptions of cyclical conditions (cost-push and demand shocks). The purpose of this section is to answer the following question: How do such noisy perceptions of the cycle affect monetary policy, inflation and the output gap? We proceed by comparing the values of those variables in the presence of imperfect information with their values under a full information benchmark. In the benchmark case policymakers possess in each period *direct information* about the realizations of the shocks up to and including the previous period by assumption. Formally, under perfect information at the beginning of period $t+1$ policy makers possess the information set J_t^* that is defined by:

$$J_t^* = \{J_t, z_{t-i}, g_{t-i}, u_{t-i} \mid i = 0, 1, 2, ...\}. \tag{3.1}$$

3.1. Consequences for monetary policy

We begin by studying the determinants of the difference between the settings of monetary policy in the presence and in the absence of the IP. Using equations (2.11), (2.21), (2.22) and (2.20), the *deviation* of the optimal interest rate in the presence of the IP from its optimal value under full information (i.e. $r^*_{t+1} = \frac{1}{\varphi}\left[\mu g_t + \frac{\lambda}{\lambda^2+(1-\beta\delta)\alpha}\rho u_t\right]$) can be written as

$$\begin{aligned}\Delta r_{t+1} &\equiv r_{t+1} - r^*_{t+1} = -\frac{1}{\varphi}\left[\mu \tilde{g}_{t|t} + \frac{\lambda\rho}{\lambda^2+(1-\beta\delta)\alpha}\tilde{u}_{t|t}\right] \\ &= \frac{\left(\mu - \rho\frac{\lambda^2}{\lambda^2+(1-\beta\delta)\alpha}\right)}{\varphi}\tilde{z}_{t|t}.\end{aligned} \tag{3.2}$$

Since $(1-\beta\delta) > 0$ it follows immediately from (3.2) that if demand shocks are sufficiently persistent in comparison to cost shocks (i.e. $\mu > \rho\frac{\lambda^2}{\lambda^2+(1-\beta\delta)\alpha}$) the deviation of the real interest rate from its full information counterpart moves in the same direction as the forecast error in potential output ($\tilde{z}_{t|t}$). Although one cannot rule out the possibility that, when the persistence in cost shocks is sufficiently larger than that of demand shocks, the opposite occurs, it appears that the first case seems more likely a-priori. The reason is that the persistence parameter of the cost shocks is multiplied by a fraction implying that Δr_{t+1} and $\tilde{z}_{t|t}$ are positively related even if ρ is larger than μ, but not by too much. Note that the smaller the (Rogoff (1985) type) conservativeness of the central bank (the higher α), the more likely it is that Δr_{t+1} and $\tilde{z}_{t|t}$ are positively related even when ρ is larger than μ. Hence, for central banks which are (using Svensson's (1997) terminology) relatively flexible inflation targeters the case in which Δr_{t+1} and $\tilde{z}_{t|t}$ are positively related is definitely the more likely one for most or all values of ρ and μ in the range between zero and one. The various possible effects of imperfect information are summarized in the following proposition:

Proposition 1. *(i) When the persistence of demand shocks is sufficiently high ($\mu > \frac{\rho\lambda^2}{\lambda^2+(1-\beta\delta)\alpha}$) monetary policy is driven mainly by "demand shocks" considerations. This implies that potential output over/under-estimation (causing the demand shock to be under/over-estimated) leads to real rates which are lower/higher than the rate which is optimal in the absence of the IP.*

(ii) When the persistence of demand shocks is sufficiently low ($\mu < \frac{\rho\lambda^2}{\lambda^2+(1-\beta\delta)\alpha}$)

monetary policy is driven mainly by "cost-push shocks" considerations. This implies that potential output over/under-estimation (causing the cost-push shock to be over/under-estimated) leads to a real rate which is higher/lower than the rate that is optimal in the absence of the IP.

To understand the intuition underlying the proposition it is useful to consider the case in which there is, in period t, a negative shock to potential output and no changes in the cyclical shocks, g and u. This leads, as of the beginning of period $t+1$, to overestimation of potential output in period t ($\tilde{z}_{t|t} < 0$). Remark 1 implies that this overestimation is associated with an overestimation of the cost shock and an underestimation of the demand shock of period t.

The policy chosen at the beginning of period $t+1$ aims to offset the (presumed) deflationary impact of the demand shock on the output gap and the (presumed) inflationary impact of the cost shock on inflation. In comparison to the full information benchmark, the first objective pushes policy towards expansionism while the second pushes it towards restrictiveness. If demand shocks are relatively persistent the first effect dominates since policymakers believe that most of what they perceive to be a negative demand shock in period t is going to persist into period $t+1$ while what they perceive to be a positive cost shock in period t is not going to persist much into period $t+1$.[9] Hence, in this case monetary policy is more expansionary than in the full information benchmark and Δr_{t+1} and $\tilde{z}_{t|t}$ are positively related (case (i) in the proposition). But if the reverse is true (cost shocks are relatively more persistent) beliefs about the cost shock in period $t+1$ dominate policy pushing it towards tightening. As a consequence monetary policy is more restrictive than in the full information benchmark and Δr_{t+1} and $\tilde{z}_{t|t}$ are negatively related (case (ii) of the proposition).

It is shown, by means of simulations, in the last part of the appendix that the basic message of proposition 1 extends to the case of backward looking models with more elaborate lag structures like that of Svensson (1997).

3.2. Consequences for the output-gap and inflation

We turn next to the consequences of mismeasurement of potential output for the output-gap and inflation. The objective is, as in the previous subsection, to

[9]This remark follows directly from the fact that $g_{t+1|t} = \mu g_{t|t}$ and $u_{t+1} = \rho u_{t|t}$.

analyze the deviations of outcomes obtained in the presence of the IP from those that arise in its absence. Using (2.12) and (2.13) it is immediate to relate these deviations to the interest rate deviations studied above. This yields:

$$\Delta x_{t+1} \equiv x_{t+1} - x^*_{t+1} = -\varphi \Delta r_{t+1} \tag{3.3}$$

$$\Delta \pi_{t+1} \equiv \pi_{t+1} - \pi^*_{t+1} = -\varphi \lambda \Delta r_{t+1} \tag{3.4}$$

where x^*_{t+1} and π^*_{t+1} are the values of the output gap and of inflation under optimal monetary policy when information is perfect. These equations show that when the interest rate is below (above) its value under perfect information both inflation and the output gap are above (below) their full information values.

The case of over-expansionary monetary policy (case (i) of proposition 1) is consistent with Orphanides (2001, 2003a) empirical results according to which, during the seventies, US monetary policy was overly expansionary due to an over-estimation of potential output and an associated underestimation of the output gap. Obviously, this underestimation could also have been due to inefficient forecasting procedures on the part of the Fed. A main message of this paper is that this effect is present even if monetary policy is ex-ante optimal and forecasting procedures are as efficient as technically feasible.[10] In normal times during which the change in potential output is not too far from its mean this effect is likely to be small and short lived. But when large permanent shocks to potential output occur this effect is likely to be large and more persistent. This point is discussed in detail in the next section.

[10] Note that, given the postulated information structure, it is possible to completely eliminate any difference between the values of the real rate, inflation and the output gap in the absence and in the presence of the IP by appropriate choice of the degree of (lack of) central bank conservativeness, α. From equations (3.2), (3.3) and (3.4) the value of α that delivers this objective is given by

$$\alpha = \frac{\lambda^2}{(1-\beta\delta)} \left(\frac{\rho}{\mu} - 1 \right).$$

But, since the preferences of society do not generally coincide with this value of α there is a tradeoff between reduction of the impact of the IP on policy and the attainment of socially desirable relative levels of stabilization of inflation and of output.

4. Optimal forecasts of potential output, serially correlated forecast errors and implications for monetary policy

This section describes the solution to the signal extraction, or filtering, problem faced by policymakers. To illustrate the basic mechanisms at work we focus in the text on the particular (but simpler) case where demand and cost push shocks are equally persistent ($\mu = \rho$), which yields a tractable closed form solution. A discussion of the procedure for obtaining the solution for the case in which the degrees of persistence differ ($\rho \neq \mu$), based on the Kalman filter, is given in Appendix B, where we show that the main qualitative properties of the optimal predictor when shocks are equally persistent carry over to the more general case.[11]

4.1. Filtering under equally persistent demand and cost-push shocks

This subsection describes the optimal predictor of potential output when demand and cost push shocks are equally persistent ($\mu = \rho$). The conditional expectation of z_t based on J_t, $z_{t|t}$, is given by (the derivation appears in Appendix A):[12]

$$z_{t|t} = aS_t + (1-a)(1-\kappa)\sum_{i=0}^{\infty}\kappa^i S_{t-1-i} \tag{4.1}$$

where :

$$\kappa \equiv \frac{2}{\phi+\sqrt{\phi^2-4}} \in (0,1) \qquad \phi \equiv \frac{2+T(1+\mu^2)}{1+\mu T} \geq 2;\ T \equiv \left(\frac{\sigma_z^2}{\sigma_g^2} + \frac{\lambda^2\sigma_z^2}{\sigma_u^2}\right)$$

$$a \equiv \frac{[(1-\mu)+(1-\kappa)+T(1-\mu\kappa)]T}{[T(1-\mu-\mu\kappa)+(1-\mu-\kappa)](1+T)+(T+\mu)(1+\mu T)} \in (0,1)$$

$$S_{t-i} \equiv s_{1,t-i} - \frac{\lambda\sigma_g^2}{\sigma_u^2+\lambda^2\sigma_g^2}s_{2,t-i} = z_{t-i} + \frac{\sigma_u^2\cdot g_{t-i} - \lambda\sigma_g^2\cdot u_{t-i}}{\sigma_u^2+\lambda^2\sigma_g^2} \tag{4.2}$$

S_{t-i} is a combined signal that summarizes all the relevant information from period's $t-i$ data. Note that it is positively related to that period's potential output and demand shocks, and negatively related to that period's cost shock. As a consequence the optimal predictor generally responds positively to current, as well as to all past, shocks to demand, and potential output, and responds negatively

[11] Although the general case eludes the possibility of an explicit analytical solution numerical simulations using an implicit analytical solution support the qualitative nature of many of the conclusions obtained explicitly in the simpler case. Details appear in section 6.

[12] This corresponds to the predictor of (the unit root) potential output, z_t, that minimizes the mean square forecast error.

to current, as well as to all past cost shocks.

The conditional forecast (4.1) possesses several key properties. First, since a and κ are both bounded between zero and one, the current optimal predictor is positively related to the current, as well as to all past combined signals. Second, the weight given to a past signal is smaller the further in the past is that signal. Third, since $a < 1$, when a positive (negative) innovation to current potential output (z_t) occurs the potential output *estimate* increases (decreases) *by less* than actual potential output. Fourth, the sum of the coefficients in the optimal predictor in (4.1) is equal to one. Finally, although the true value of potential output is contained only in the signals $s_{1,t-i}$, the optimal predictor also assigns positive weights to the signals $s_{2,t-i}$. The intuitive reason is that, by allowing a more precise evaluation of the demand shock, g_t, the utilization of $s_{2,t-i}$ facilitates the separation of g_t from z_t in the signals $s_{1,t-i}$.

4.2. Optimal learning produces serial correlation in forecast errors of potential output and of the output gap

The form of the optimal predictor in (4.1), in conjunction with the fact that all coefficients are positive and sum to one implies that when a single shock to potential output occurs (say) in period t and persists forever without any further shocks to potential output, policymakers do not recognize its full impact immediately. Although their forecasting is optimal policymakers learn about the permanent change in potential output gradually. Initially (in period $t + 1$) they adjust their perception of potential output by the fraction a. In period $t + 2$ they internalize the larger fraction $a + (1 - a)(1 - \kappa)$, in period $t + 3$ they internalize the, even larger, fraction $a + (1 - a)(1 - \kappa) + (1 - a)(1 - \kappa)\kappa$, and so on. After many periods this fraction tends to 1, implying that after a sufficiently large number of periods the full size of the shock is ultimately learned. Thus, equation (4.1) implies that there is gradual learning about potential output and that forecast errors are, therefore, on the same side of zero during this process.

Conversely, when a single relatively large shock to the cyclical component of demand occurs it is partially interpreted for some time as a change in potential output. This too creates ex-post serial correlation in errors of forecast in the output gap and in potential output. In general two kinds of errors can be made. A change in potential output may be partly misinterpreted as a cyclical change, or a cyclical change may be partly misinterpreted as a change in potential output. Both errors tend to create ex-post serial correlation in forecast errors, but this serial correlation cannot be utilized in real time to improve policy because, unlike forecast errors of variables that become known with certainty one period after their realization, potential output of period t is not known with certainty even after that period. As a consequence the forecast error committed in period t cannot be used to "correct" future forecasts of potential output in the same way that errors in the forecast of a variable that is revealed one period after the formation of that forecast are normally used to update future forecasts.[13]

It can be shown that forecast errors of potential output and the output gap are generally serially correlated even in the population. The remainder of this subsection establishes this fact more precisely and identifies conditions under which this serial correlation is dominated by the variability of innovations to potential output. Note first, from equation (2.20), that the error in forecasting the output gap is equal to the negative of the error in forecasting potential output. Hence, if forecast errors of potential output are serially correlated, so are forecast errors of the output gap. It is shown in Appendix B that the covariance between two adjacent forecast errors is given by

$$\begin{aligned} E\left[\tilde{z}_{t|t}.\tilde{z}_{t-1|t-1}\right] &= \frac{(1-a)^2\kappa}{1-\kappa^2}\sigma_z^2 + \\ & \left(\tfrac{\sigma_u^2}{\sigma_u^2+\lambda^2\sigma_g^2}\right)^2 \left\{ \begin{array}{c} a(\mu a+\theta)+(\mu a+\theta)(\mu^2 a+\mu\theta+\theta\kappa)+ \\ (\mu^2 a+\mu\theta+\theta\kappa)(\mu^3 a+\mu^2\theta+\mu\theta\kappa+\theta\kappa^2)+.. \end{array} \right\} \sigma_g^2 \\ & + \left(\tfrac{\lambda\sigma_g^2}{\sigma_u^2+\lambda^2\sigma_g^2}\right)^2 \left\{ \begin{array}{c} a(\rho a+\theta)+(\rho a+\theta)(\rho^2 a+\rho\theta+\theta\kappa)+ \\ (\rho^2 a+\rho\theta+\theta\kappa)(\rho^3 a+\rho^2\theta+\rho\theta\kappa+\theta\kappa^2)+.. \end{array} \right\} \sigma_u^2 \end{aligned} \tag{4.3}$$

where

$$\theta \equiv (1-a)(1-\kappa). \tag{4.4}$$

[13] When the true value of the variable that is being forecasted is revealed with certainty with a lag of one period, as is often assumed, the general principle that forecast errors are serially uncorrelated in the population applies. This feature has been used extensively to test for the efficiency of financial market. However when, as is the case here, the true value of the variable that is being forecasted is not revealed with certainty even after the fact, forecast errors are serially correlated in general.

Since, except for the extreme case in which $a = 0$ and $\kappa = 1$ all terms on the right hand side of equation (4.3) are positive, errors in forecasting potential output exhibit a positive serial correlation.[14] This leads to the following

Proposition 2. *Errors in forecasting potential output and the output gap generally display a positive serial correlation.*

Interestingly this proposition is consistent with recent empirical findings in Orphanides (2003b). Orphanides utilizes real time data on the perceptions of policymakers about potential output during the 1970's and compares those perceptions with current estimates (as of October 1999) of the historical data. Taking the "current" rendition of estimates of potential output as a proxy for the true values of potential output during the seventies he finds highly persistent deviations between the current and the real time estimates of the output gap (see his Figure 3 in particular).

4.3. The deeper origins of serial correlation in forecast errors

Examination of equation (4.3) reveals that this positive serial correlation is generally due to persistence in both potential output and in the two cyclical components of the economy. The following discussion identifies conditions on the underlying variances of the innovations to potential output and to demand and costs under which this serial correlation is due mainly to shocks to potential output, and conditions under which it is due mainly to shocks to the cyclical components. In particular we will focus on the relative sizes of the variances of shocks to potential output and to the cyclical components of the economy. As a prelude to the main discussion of those issues we note the following properties of the optimal predictor

Remark 2. *(i) The coefficient, a, of the most recent observation on the compound signal in equation (4.1) is a monotonically increasing function of the ratios of variances σ_z^2/σ_g^2 and σ_z^2/σ_u^2. When both of those ratios tend to zero, a tends to zero too, and when both of them tend to infinity, a tends to one.*

[14] As a matter of fact this property is recognized by the time-series literature on unobserved component models. For instance, the prediction errors produced by the Kalman filter about a variable that is not-observable, even on an expost basis, have a non-zero covariance matrix in the population (equation 13.2.27 in Hamilton (1994) can be used to compute this covariance).

(ii) The combination of parameters, κ, in equation (4.1) is a monotonically decreasing function of the ratios of variances σ_z^2/σ_g^2 and σ_z^2/σ_u^2. When both of those ratios tend to zero κ tends to one.

The proof appears in Appendix C. An immediate implication of this Remark is that, when the variance, σ_z^2, of innovations to potential output is relatively small, a is not far from zero and $(1-a)$ and κ are not far from one, implying that θ in equation (4.4) is not far from zero. But inspection of equation (4.3) reveals that when a and θ are not far from zero the coefficients of σ_g^2 and of σ_u^2 in equation (4.3) are nearly zero while (since κ is not far from one) the coefficient of σ_z^2 is rather large. As σ_z^2 rises the coefficients of σ_g^2 and of σ_u^2 increase and the coefficient of σ_z^2 decrease.

Since as σ_z^2 goes up its coefficient goes down, it would appear that the effects of an increase in σ_z^2 on the size of the contribution of shocks to potential output to the serial correlation in forecast errors of potential output is ambiguous. Although this ambiguity may apply for values of σ_z^2 above a certain threshold, it does not hold for small values of σ_z^2. The reason is that, for small values of σ_z^2, the size of the derivative of the product $\frac{(1-a)^2\kappa}{1-\kappa^2}\sigma_z^2$ with respect to σ_z^2 is dominated by the term $\frac{1}{1-\kappa^2}$ which is positive and large relative to all the other components of this derivative since the denominator in this expression is very small. This observation, in conjunction with the fact (implied by Remark 2) that the derivative of κ with respect to σ_z^2 is negative, implies that, below some threshold, the lower the variability of innovations to potential output, the higher the contribution of this variability to the serial correlation in forecast errors.

Those observations are summarized in the following proposition.

Proposition 3. *(i) When σ_z^2 is sufficiently low the serial correlation in forecast errors of potential output and of the output gap is caused mainly by innovations to potential output while the effect of innovations to demand and costs on this serial correlation is negligible.*

(ii) At the other extreme, when σ_z^2 is sufficiently large in comparison to σ_g^2 and σ_u^2, $(1-a)$ tends to zero and the serial correlation in forecast errors of potential output and of the output gap is caused mainly by innovations to demand and costs while the effect of innovations to potential output on this serial correlation is negligible.

An implication of the proposition is that when the variability of innovations to potential output is small the, relatively rare, occurrence of a large shock to potential output will induce a large and sustained sequence of serially correlated errors. Since the innovation to potential output is relatively large and since learning is highly gradual in this case, the shock dominates the learning process for some time. As a consequence when looking backwards, forecast errors in potential output and the resulting monetary policy "errors" will be serially correlated. The intuitive reason is that the shock to potential output is partially interpreted for quite some time as a persistent change in the output gap. Simulations presented in the last part of the appendix show that the basic message of propositions 2 and 3 extend to the case of backward looking models with more elaborate lag structures.

4.4. Implications for monetary policy during the seventies and the nineties

Proposition 2 states that there is always serial correlation in the population, but the first part of proposition 3 implies that this serial correlation is particularly in evidence following the realization of a large change in potential output. The reason is that, in finite samples, the magnitude of the serial correlation is directly related to the size of the shock to potential output.[15] This view implies that the economic events of the seventies can be viewed as having been triggered by a large decrease in potential output about which policymakers learned quite gradually but optimally. This point of view fits surprisingly well the persistent downward revisions in forecasted potential output during the end of the seventies, as can be seen by taking a second look at Figure 1 in the introduction.

The main lessons from these remarks are summarized in the following proposition.

Proposition 4. *When σ_z^2 is small, optimal monetary policy in the aftermath of a period characterized by the realization of a large permanent change to potential output appears ex-post as being systematically biased in a certain direction for some time.*

[15] Cukierman and Meltzer (1982) use this feature to show (in the context of tests of efficiency in financial markets) that this mechanism will produce serially correlated forecast errors in finite samples even when there is no serial correlation in the population.

(i) When the potential output shock is negative policy is too loose in comparison to the full information benchmark. Although optimal in "real time", this policy stance is retrospectively judged as being too loose.

(ii) When the potential output shock is positive policy is too restrictive in comparison to the full information benchmark. In particular, a large increase in potential output induces policymakers to behave in a way that overemphasizes the concern for price stability. Although optimal in "real time", this policy stance is retrospectively judged as being too restrictive.

The first part of the proposition corresponds to the retrospectively loose monetary policy of the seventies identified by Orphanides (2001, 2003a). This retrospective policy error was triggered by overestimation of potential output and underestimation of the output gap. The second part of the proposition appears to fit the "new economy" of the nineties. The large positive technological shock to potential output during the nineties was initially interpreted partly as a positive output gap and triggered a policy response that was judged retrospectively to be overly restrictive.

5. Implications of other differences between the seventies and the nineties

Taken literally, the previous analysis implies that, other things remaining the same and except for the sign of policy errors, the seventies and the nineties are similar. In the seventies monetary policy was too loose in comparison to a perfect information benchmark because potential output was overestimated and in the nineties it was overly restrictive because, at least initially, potential output was underestimated. But other things did not remain the same between those two periods. In particular, there is reason to believe that at least two other things changed between the seventies and the nineties.

First the relative emphasis of policy on price stabilization versus output stabilization shifted towards the former. In terms of our model this means that the parameter α decreased between the seventies and the nineties implying, via equation (2.11), that the response of the interest rate to cost shocks in the nineties was stronger than in the seventies. Arguments and evidence presented in Taylor

(1998), Clarida, Galí and Gertler (2000) and Siklos (2002, pp. 61-64) are consistent with this view.[16] Second, during the seventies policymakers might have been overly optimistic about their ability to forecast potential output and the natural level of employment. In what follows we use the analytical framework of the paper to investigate the consequences of those two changes for the comparison between the seventies and the nineties.

5.1. Consequences of changes in central bank conservativeness between the seventies and the nineties

Proposition 1 implies that, provided $\mu > \frac{\rho\lambda^2}{(1-\beta\delta)\alpha+\lambda^2}$, overestimation of potential output ($\tilde{z}_{t|t} < 0$) leads to real rates that appear, with the benefit of hindsight, to have been too low. Assuming that this condition was satisfied during the seventies, it follows that, for a given absolute value of the forecast error ($| \tilde{z}_{t|t} |$) the absolute deviation of the interest rate from its full information benchmark is proportional to the difference $\mu - \rho\lambda^2/((1-\beta\delta)\alpha + \lambda^2)$. Since during the nineties central banks were more conservative, $\alpha_{70s} > \alpha_{90s}$, this implies that

$$\mu - \rho\lambda^2/((1-\beta\delta)\alpha_{70s} + \lambda^2) > \mu - \rho\lambda^2/((1-\beta\delta)\alpha_{90s} + \lambda^2) > 0.$$

This leads to:

Proposition 5. *For a given absolute value of the forecast error in potential output,* $| \tilde{z}_{t|t} |$, *retrospective policy errors are larger during the seventies than during the nineties.*

The proposition implies that even if the standard deviation of the shocks to potential output was similar during the seventies and during the nineties, policy errors

[16] Actually those studies pertain to estimates of the relative weight on the inflation gap in the central bank's policy rule rather than to the relative weight ($1/\alpha$) on inflation in the central bank's loss function. Obviously the weight in the reaction function could have gone up over the past several decades, due to changes in the structure of the economy, even if the loss function weight had remained constant.

However it is likely that a non negligible part of the increase in the former is due to an increase in the loss function weight. The reason is that there is direct evidence suggesting that the level of central bank independence (or conservativeness) has been increasing, world wide over the last several decades (Cukierman (1998) and Cukierman, Miller and Neyapti (2002)). It is likely that this worldwide process also raised the effective level of conservativeness of the Fed following the seventies.

should prove to have been smaller in the second period. The intuitive reason is that the increased focus on the stabilization of inflation between the two periods reduced the divergence between optimal policy under imperfect and under full information about potential output and about the cyclical shocks, g_t and u_t. The discussion in Taylor (1998) and casual observation appear to be consistent with this implication of the analysis. More generally the analysis suggests that, in the presence of uncertainty about potential output, central bank conservativeness affects the economy not only directly (as in Rogoff (1985) or Walsh (1995)) but also through the signal extraction problem solved by policymakers.

5.2. Consequences of an increase in awareness about uncertainty with respect to potential output

During the sixties and the early seventies policymakers were generally confident about their ability to control the economy implying that they most likely had a good opinion of their forecasting ability. The high hopes generated by the construction of large scale econometric models like the FMP model attest to that. After the oil shocks of the seventies and the associated great inflation policymakers became more aware of the limits of their own forecasting ability.

We embed this presumption into the model by assuming that during the seventies policymakers were overoptimistic about potential output uncertainty into the analysis by postulating that during the seventies the perceived variance, σ^2_{zp}, of the innovation to potential output was lower than the true variance, σ^2_z, but that during the nineties the perceived variance adjusted upwards and became equal to the true variance. For the rest, we maintain the hypothesis that the stochastic processes generating potential output and the cyclical shocks remained the same over the entire period, and that, given the perceived variance in each period, policymakers used optimal filters and chose policy so as to minimize expected losses. This is a stylized way to isolate the consequences of overconfidence about estimates of potential output during the seventies. An immediate consequence of these presumptions is that the mean square error in forecasting potential output during the seventies was larger than the optimal mean square error.[17] By contrast,

[17] This is a direct consequence of the presumption that, although they used the correct form for the predictor, policymakers during the seventies fed this predictor with the lower perceived variance, σ^2_{zp}, rather than with the actual variance, σ^2_z.

during the nineties those two forecast errors were equal.

Before continuing we digress to the following proposition

Proposition 6. *For the case $\mu = \rho$, the higher σ_z^2, the higher the relative size of the weights on more recent observations of the combined signal, S_{t-i}, in equation (4.1), and the lower the relative size of the weights on relatively distant past observations on S_{t-i}.*

The proof is obtained by differentiating the parameters κ and a in equation (4.1) with respect to σ_z^2, by showing that κ is a decreasing function of σ_z^2 and that a is an increasing function of σ_z^2 and by noting that the sum of the weights on the combined signal is equal to one for **all** values of σ_z^2.

Together with the presumption that during the seventies $\sigma_{zp}^2 < \sigma_z^2$ while during the nineties $\sigma_{zp}^2 = \sigma_z^2$, the proposition implies that, in addition to being more accurate on average, learning about changes in potential output during the nineties was quicker than in the seventies. On this view monetary policy during the nineties was nearer to its full information optimal value in comparison to the seventies thanks in part to a swifter and more accurate recognition of changes in potential output.

6. Concluding remarks

This paper provides a unified explanation to account for part of the inflation of the seventies and for part of the remarkable price stability of the nineties. This is accomplished by showing that, even if monetary policy is optimal and forecasts of potential output are efficient, large permanent changes in potential output trigger excessively loose monetary policy when those changes are negative and excessively tight policy when the changes are positive. But the paper also shows that even if the positive shocks to potential output during the nineties were similar in absolute value to the negative shocks of the seventies, there is reason to believe that policy was excessively loose in the seventies to a greater extent than it was excessively tight during the nineties. This conclusion is based on two presumptions. The first is that the Fed was relatively more conservative in the Rogoff (1985) sense in the nineties than in the seventies. For the economic structure postulated in the paper, a higher degree of conservativeness reduces the difference between the imperfect and the full information policy at any given level of the error in forecasting potential output. The second is that a more realistic evaluation of uncertainties surrounding potential output enabled the Fed to learn faster and more accurately about changes in potential output during the nineties than during the seventies, so that its policy was nearer to the full information benchmark.

The framework of the paper also leads to two wider conclusions that are likely to transcend the particular model used to illustrate them. The first is that even if monetary policy is chosen optimally and even if, given the stochastic structure of shocks, available information is used as efficiently as possible, retrospective policy errors are unavoidable. During periods in which changes in potential output are moderate these errors are neither very important, nor persistent. As a consequence, they do not draw much attention ex-post. But during periods following large sustained changes in potential output, retrospective policy errors appear, with the benefit of hindsight, to be large and to exihibit substantial serial correlation. This makes them noticeable and draws public attention. Thus, even central banks that forecast and behave optimally may sometimes be judged retrospectively as having committed serious policy errors. But, since they had behaved efficiently at the time, it does not follow that (given the information structure) such errors can be avoided in the future. This mechanism is quantitatively more important the smaller the relative size of the variance of innovations to potential output.

Obviously, this does not necessarily mean that policy and forecasting procedures during the seventies were as efficient as possible at the time. The point, however, is that the ex-post identification of policy errors is not sufficient to conclude that such errors were avoidable in real time. A challenge facing policymakers and economists is to distinguish between avoidable (in real time) and unavoidable policy errors. We believe that models in the spirit of the one analyzed here, where policy is consistent with the economic structure and information is processed efficiently, can pave the way towards a better understanding of this issue.

The second conclusion is that, with the exception of extreme cases, the fact that in the wake of large and sustained changes in potential output policymakers commit serious errors in forecasting potential output does not imply that noisy but optimally devised forecasts of potential output should not be used as indicator variables for monetary policy.

In order to focus on the consequences of imperfect information in isolation we have used a backward looking economc structure that abstracts from the effects of expectations. Simulations (not shown) we conducted with the Ehrmann and Smets (2003) model of the Euro area which features both forward as well as backward looking terms, suggest that some of our results carry over to this more elaborate framework while others are modified. In particular, retrospective errors in the conduct of monetary policy still exihibit serial correlation and, following a positive shock to potential output, the interest rate overshoots its full information counterepart for a while. But, after a while the interest rate undershoots the full information benchmark. In addition, the simulations suggest that the effects of imperfect information do not differ by much between a discretionary and a commitment regime.

We do not know the extent to which these new results depend on the particular paramaters in the Ehrmann and Smets model rather than on the introduction of expectations per se. Discrimination between those two views could be helped by further analytical work on an imperfect information model that features both forward as well as backward looking terms. This is left for future work.

A. Appendix: The Filtering Problem

At time $t+1$ the policy maker's problem is to estimate z_t based on $J_{t,}$ i.e. using all the information contained in the observed sequence of signals $s_{1,t-i}$ and $s_{2,t-i}$ ($i = 0, 1, 2, ...$). To this end, it is convenient to define the new signal $s_{3,t-i} \equiv s_{1,t-i} - \frac{1}{\lambda} s_{2,t-i}$. Rewriting the linear predictor for z_t conditional on J_t as:

$$\begin{aligned} z_{t|t} &\equiv \sum_{i=0}^{\infty} a_i \cdot s_{1,t-i} + \sum_{i=0}^{\infty} b_i \cdot s_{3,t-i} \\ \text{where } s_1 &= z_t + g_t \text{ and } s_{3,t} = z_t - (1/\lambda)u_t \end{aligned} \tag{A.1}$$

and the last line follows directly from (2.15) and (2.16). We seek to determine optimal weights a_i and b_i that minimize the mean square forecast error of the z_t predictor (it follows from this property that the predictor z_t^* equals the expectation of z_t conditional on J_t i.e. $z_{t|t}$). This amounts to solving $\min\limits_{a_i,b_i} Q$,where:

$$\begin{aligned} Q &\equiv E\left\{\left[z_t - z_{t|t}\right]^2 \mid J_t\right\} = \\ &= \sigma_z^2 \left\{ \begin{array}{c} [1-(a_0+b_0)]^2 + [1-(a_0+b_0)-(a_1+b_1)]^2 + .. \\ .... + [1-(a_0+b_0)-(a_1+b_1) - ... - (a_i+b_i)]^2 + ... \end{array} \right\} + \\ &\quad + \sigma_g^2[(a_0^2 + (\mu a_0 + a_1)^2 + (\mu^2 a_0 + \mu a_1 + a_2)^2 + ... + (\mu^i a_0 + ... + a_i)^2 + ...] + \\ &\quad + \frac{\sigma_u^2}{\lambda^2}[(b_0^2 + (\rho b_0 + b_1)^2 + (\rho^2 b_0 + \rho b_1 + b_2)^2 + ... + (\rho^i b_0 + ... + b_i)^2 + ...] \end{aligned} \tag{A.2}$$

The first order conditions with respect to a_i and b_i, for $i = 0, 1, 2, ..$yield respectively:

$$\begin{aligned} 0 &= -\sigma_z^2 \left\{ \begin{array}{c} [1-(a_0+b_0) - .. - (a_i+b_i)] + \\ +[1-(a_0+b_0) - ... - (a_i+b_i) - (a_{i+1}+b_{i+1})] + ... \end{array} \right\} + \\ &\quad + \sigma_g^2 \left[(\mu^i a_0 + ... + a_i) + \mu(\mu^{i+1} a_0 + ... + a_{i+1}) + \mu^2(\mu^{i+2} a_0 + ... + a_{i+2}) + ..\right] \end{aligned} \tag{A.3}$$

and

$$\begin{aligned} 0 &= -\sigma_z^2 \left\{ \begin{array}{c} [1-(a_0+b_0) - .. - (a_i+b_i)] + \\ +[1-(a_0+b_0) - ... - (a_i+b_i) - (a_{i+1}+b_{i+1})] + ... \end{array} \right\} + \\ &\quad + \frac{\sigma_u^2}{\lambda^2} \left[(\rho^i b_0 + ... + b_i) + \rho(\rho^{i+1} b_0 + ... + b_{i+1}) + \rho^2(\rho^{i+2} b_0 + ... + b_{i+2}) + ..\right]. \end{aligned} \tag{A.4}$$

Note that the two first order conditions (FOC) have an identical first term inside the curly bracket and a similar form for the term in the second curly bracket, which only differ in that μ (a_i) is replaced by ρ (b_i). Leading (A.3) by one step, multiplying the resulting expression by μ and subtracting it from (A.3) yields:

$$\begin{aligned} 0 &= -\sigma_z^2 \left\{ \begin{array}{c} [1-(a_0+b_0)-..-(a_i+b_i)]+ \\ +(1-\mu)\left[(1-(a_0+b_0)-...-(a_i+b_i)-(a_{i+1}+b_{i+1}))+....\right] \end{array} \right\} + \\ & +\sigma_g^2(\mu^i a_0+...+a_i) \qquad \text{(A.5)} \end{aligned}$$

Leading (A.5) by one step and subtracting the resulting expression from (A.5) yields

$$\begin{aligned} 0 &= -\sigma_z^2 \left\{\mu(a_{i+1}+b_{i+1})+(1-\mu)\left[(1-(a_0+b_0)-...-(a_i+b_i))\right]\right\}+ \\ & +\sigma_g^2\left[(1-\mu)(\mu^i a_0+...+a_i)-a_{i+1}\right] \qquad \text{(A.6)} \end{aligned}$$

Leading (A.6) by one step and subtracting the resulting expression from (A.6) yields

$$\begin{aligned} 0 &= -\sigma_z^2\left[(a_{i+1}+b_{i+1})-\mu(a_{i+2}+b_{i+2})\right]+ \qquad \text{(A.7)} \\ & +\sigma_g^2\left[(1-\mu)^2(\mu^i a_0+...+a_i)-(2-\mu)a_{i+1}+a_{i+2}\right] \end{aligned}$$

Leading (A.7) by one step, multiplying the resulting expression by $1/\mu$ and subtracting it from (A.7) yields

$$0=\sigma_z^2\left[(a_i+b_i)\mu-(a_{i+1}+b_{i+1})(1+\mu^2)+(a_{i+2}+b_{i+2})\mu\right]+\sigma_g^2\left[a_i-2a_{i+1}+a_{i+2}\right] \qquad \text{(A.8)}$$

Applying to the FOC for b_i (A.4) algebraic transformations identical to those used to establish (A.8) leads to

$$0=\sigma_z^2\left[(a_i+b_i)\rho-(a_{i+1}+b_{i+1})(1+\rho^2)+(a_{i+2}+b_{i+2})\rho\right]+\frac{\sigma_u^2}{\lambda^2}\left[b_i-2b_{i+1}+b_{i+2}\right] \qquad \text{(A.9)}$$

where both (A.8) and (A.9) hold for $i=0,1,2,3,....$ These two equations constitute a system of two homogenous linear second order difference equations in the unknowns a_i and b_i. We next solve explicitly for the simpler case in which $\mu=\rho$ and then present the general solution using the Kalman filter.

A.1. The case of equally persistent demand and cost-push shocks

When $\mu=\rho$ the difference equations (A.8) and (A.9) can be uncoupled. Since the right hand sides of those equations are equal for all $i's$ the a_i and b_i are related by the linear relationship

$$b_i = a_i \frac{\lambda^2 \sigma_g^2}{\sigma_u^2} \qquad \text{for } i = 0, 1, 2, ... \tag{A.10}$$

where the equality for $i = 0$ is established from the first order conditions for a_0 and b_0. Substituting this expression for b_i into (A.8) yields

$$\begin{aligned} 0 &= a_i - \phi a_{i+1} + a_{i+2} \qquad \text{for } i = 1, 2, ... \\ \text{where } \phi &\equiv \frac{2 + T(1+\mu^2)}{1 + T\mu} \qquad \text{and } T \equiv \left(\frac{\sigma_z^2}{\sigma_g^2} + \frac{\lambda^2 \sigma_z^2}{\sigma_u^2} \right) \end{aligned} \tag{A.11}$$

Equation (A.11) has one non-explosive solution which is given by

$$a_i = a_1 \kappa^{i-1} \qquad \text{for } i = 1, 2, ... \tag{A.12}$$

where a_1 is a constant term to be determined and κ is the "stable" root (i.e. smaller than one) of the second order equation in κ: $\kappa^2 - \phi\kappa + 1$ (the characteristic equation corresponding to A.11). It remains to determine the values of a_0 and of a_1. Equation (A.6) for $i = 0$ can be re-arranged (using (A.10) to get rid of b_i) to yield:

$$a_1 \equiv \frac{(1-\mu)\left[(1+T)a_0 - \frac{\sigma_z^2}{\sigma_g^2}\right]}{(1+\mu T)}. \tag{A.13}$$

Equation (A.6) for $i = 1$ can be re-arranged (using $a_2 = a_1 \kappa$ to get rid of a_2) to yield:

$$a_1 \equiv \frac{(1-\mu)\left[\frac{\sigma_z^2}{\sigma_g^2} - (T+\mu)a_0\right]}{T(1-\mu-\mu\kappa) + (1-\mu-\kappa)}. \tag{A.14}$$

The solutions for a_0 and a_1 are determined by the system: (A.13), (A.14). Using (A.10), (A.12) and the expression for the optimal predictor in (A.1) the conditional expectation of z_t can thus be written as

$$
\begin{aligned}
z_{t|t} &= a_0 S_t^{'} + a_1 \sum_{i=0}^{\infty} \kappa^i S_{t-1-i}^{'} &\quad \text{(A.15)}\\
\text{where} &: \\
a_0 &\equiv \frac{[(1-\mu)+(1-\kappa)+T(1-\mu\kappa)]\frac{\sigma_z^2}{\sigma_g^2}}{[T(1-\mu-\mu\kappa)+(1-\mu-\kappa)](1+T)+(T+\mu)(1+\mu T)} \in (0, 1+\frac{\lambda^2\sigma_g^2}{\sigma_u^2}) \\
a_1 &\equiv \frac{(1-\mu)(1+T)a_0 - \frac{\sigma_z^2}{\sigma_g^2}(1-\mu)}{(1+\mu T)} \\
S_{t-i}^{'} &\equiv s_{1,t-i} + \frac{\lambda^2\sigma_g^2}{\sigma_u^2}(s_{1,t-i} - \frac{1}{\lambda}s_{2,t-i}) = \left(1+\frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right) z_t + g_{t-i} - \frac{\lambda\sigma_g^2}{\sigma_u^2} u_{t-i}
\end{aligned}
$$

Some algebra reveals that

$$
a_0 + \frac{a_1}{1-\kappa} = \left(1+\frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right) \tag{A.16}
$$

suggesting the convenient reformulation of the filter used in the main text, that is based on the modified signal $S_{t-i} \equiv S_{t-i}^{'}\left(1+\frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)^{-1}$. Defining $a \equiv a_0\left(1+\frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)^{-1}$ and $a_1^{'} \equiv a_1\left(1+\frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)^{-1}$ and noting that (A.16) in conjunction with those definitions imply $a + a_1^{'}/(1-\kappa) = 1$, it follows that $a_1^{'} = (1-\kappa)(1-a)$. The optimal predictor in the main text, (4.1), follows by using those relations in (A.15).

A.2. Solution for the general case ($\mu \neq \rho$) using the Kalman filter

When $\mu \neq \rho$ the second-order difference equations system given by (A.8) and (A.9) can not be uncoupled and computing a closed-form analytical solution for the optimal filter is more involved. In the following we solve the filtering problem by applying the Kalman filter. We begin by rewriting the system of equations (2.3), (2.4) and (2.5) in matrix form as

$$
x_{t+1} = Ax_t + Cw_{t+1} \tag{A.17}
$$

where

$$
x_{t+1} \equiv \begin{bmatrix} z_{t+1} \\ g_{t+1} \\ u_{t+1} \end{bmatrix}, A \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \rho \end{bmatrix}, C \equiv \begin{bmatrix} \sigma_z & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_u \end{bmatrix}, \tag{A.18}
$$

and where w_{t+1} is a vector of iid innovation with unit variance. The system in equation (A.17) is the Kalman filter's state equation. Rewriting equations (2.15) and (2.16) in matrix form we obtain

$$y_t = Gx_t + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{A.19}$$

where

$$y_t \equiv \begin{bmatrix} s_{1,t} \\ s_{2,t} \end{bmatrix}, \; G \equiv \begin{bmatrix} 1 & 1 & 0 \\ 0 & \lambda & 1 \end{bmatrix}. \tag{A.20}$$

Equation (A.19) is the measurement equation of the Kalman filter for our system. A general specification of the state and measurement equations is given by equation (8.1) in chapter 8 of Hansen and Sargent (1997). Equations (A.17) and (A.19) correspond, for our system, to equation (8.1) of that chapter.[18] Algebraic manipulation of equations (8.8) and (8.9) in conjunction with equation (8.11) of that chapter imply that, for the case in which the covariance matrix Σ of the one-step ahead forecast error in the state variables (i.e. $x_t - x_{t|t-1}$) has converged, the optimal forecasts of the hidden states in x_t, given the information set J_t, are given by

$$x_{t|t} = x_{t|t-1} + K\left[y_t - Gx_{t|t-1}\right] \tag{A.21}$$

where

$$K \equiv \Sigma G' \left[G\Sigma G'\right]^{-1} \tag{A.22}$$

and

$$\Sigma = A\Sigma A' + CC' - A\Sigma G' \left[G\Sigma G'\right]^{-1} G\Sigma A'. \tag{A.23}$$

Equation (A.23) implicitly determines the unknown matrix, Σ, and given Σ, equation (A.22) determines K. Equation (A.21) can be rewritten as

$$x_{t|t} = \left[I - KG\right] x_{t|t-1} + Ky_t. \tag{A.24}$$

[18]Since there is no measurement error in our system the variance - covariance matrix of the noise in the measurement equation is identically zero. There is nonetheless a meaningful signal extraction problem because there are only two signals and three hidden states.

Table A.1: Baseline parameter values

Parameters			Innovations		
μ	ρ	λ	σ_z	σ_u	σ_g
.6	.5	.05	(.01; .30)	.15	.10

Lagging (A.24) by one period and using $x_{t+1|t} = Ax_{t|t}$, repeated substitution of the resulting expression into (A.24) yields

$$x_{t|t} = \sum_{j=0}^{\infty} D^j K y_{t-j} \tag{A.25}$$

where

$$D \equiv [I - KG]\, A,\ D^0 \equiv I \tag{A.26}$$

and D^j is the $j - th$ power of D. Note that the matrix $D^j K$ is of order 3 by 2. Denoting by k_{11}^j and k_{12}^j the first and second elements in the first row of $D^j K$ and using equation (A.25), the optimal predictor of potential output can be written as

$$z_{t|t} = \sum_{j=0}^{\infty} k_{11}^j S_{t-j} \tag{A.27}$$

where

$$S_{t-j} \equiv s_{1,t-j} + \omega^j \cdot s_{2,t-j},\ j = 0, 1, ...\infty. \tag{A.28}$$

$$\omega^j \equiv \frac{k_{12}^j}{k_{11}^j} \tag{A.29}$$

Solving for the optimal filter numerically using Matlab reveals that the key properties of the predictor that were established analytically in the case $\mu = \rho$ are preserved in the more general case. Table B1 reports the benchmark parametrization of one such example. Since a key variable in the signal extraction process is the relative size of the innovations to potential output versus those in g and u, we let the standard deviation of potential output σ_z vary between .01 and .3 to show how the properties of the optimal filter vary as the signal to noise ratio in the fundamentals changes.

The experiments show the following. (i) The sum of the coeff $\sum_{j=0}^{\infty} k_{11}^j = 1$ (ii)

The coefficients k_{11}^j are decreasing in j, i.e. the weight attributed to the observable S_t gets smaller as S_t gets older. Figure A1 plots the coefficients k_{11}^j for the first six lags ($j = 0, 1, .., 5$) computed from the optimal filter for four different values of σ_z (ranging from relatively small, $\sigma_z = 0.01$, to relatively large, $\sigma_z = 0.31$).

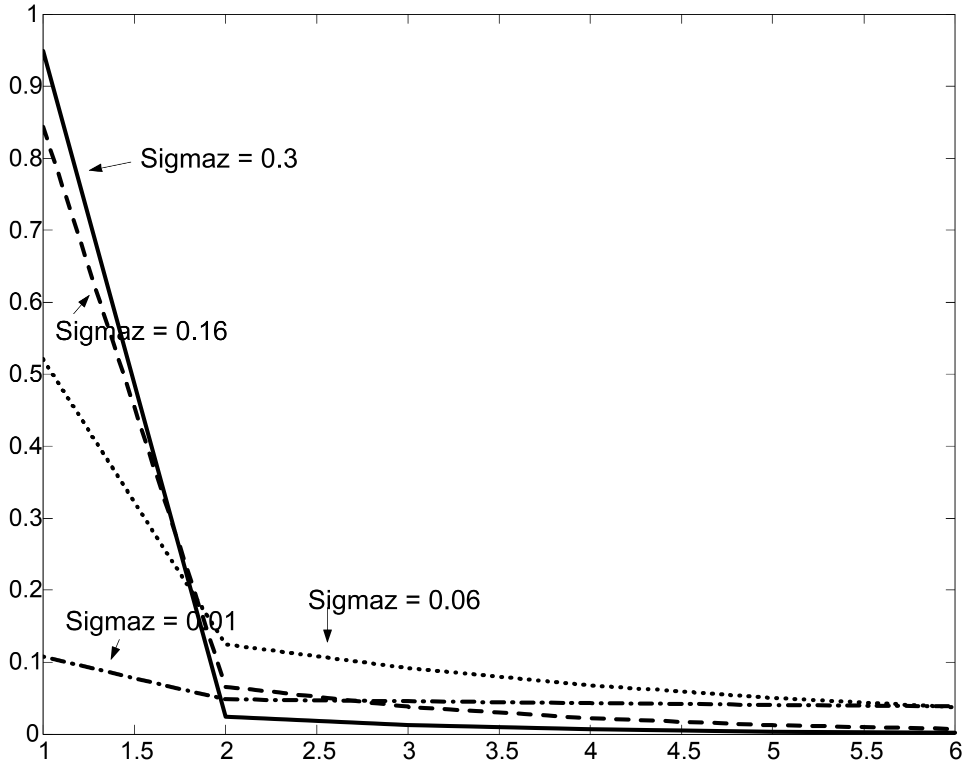

Figure A1: Weights k_{11}^j on Observables for $j = 0, 1, .., 5$.

The decreasing profile of each of the four curves in the figure indicates that the value of the information contained in the observable, S_t, decreases as that observation ages. The magnitude of the innovation in z , σ_z, relative to the size of the other innovations in the system (σ_u and σ_g) is a key determinant of the speed at which the value of information "depreciates". As this relative volatility increases, the observables contain a better signal about z and the value of past observation therefore diminishes. This is apparent from the figure where, as σ_z increases, the weight on the current signal grows (from around 0.1 to above 0.9 in our example); since the sum of all the k_{11}^j weights is 1, an increase in k_{11}^0implies that the sum of the remaining coefficients, i.e. the weight attached to past observables, decreases as σ_z increases.

B. Appendix: The serial correlation properties of errors in forecasting potential output

Rewriting the optimal predictor in equation (4.1) as $z_{t|t} = \Sigma_{i=0}^{\infty} d_i S_{t-i}$ where $d_0 \equiv a$ and $d_i \equiv (1-a)(1-\kappa)\kappa^{i-1}$ for $i \geq 1$, substituting this form of the predictor into the expression for the forecast error in equation (2.19) and regrouping terms so as to express this error in terms of infinite sums of the innovations in z, g and u we obtain

$$\tilde{z}_{t|t} \equiv Z_t - \frac{\sigma_u^2}{\sigma_u^2 + \lambda^2 \sigma_g^2} G_t + \frac{\lambda \sigma_g^2}{\sigma_u^2 + \lambda^2 \sigma_g^2} U_t \tag{B.1}$$

where

$$\begin{aligned} Z_t &\equiv \sum_{i=1}^{\infty} d_i \left[\hat{z}_{t-1} + .. + \hat{z}_{t-i} \right] \\ G_t &\equiv \sum_{i=0}^{\infty} d_i \left[\hat{g}_{t-i} + \mu \hat{g}_{t-i-1} + \mu^2 \hat{g}_{t-i-2} + .. \right] \\ U_t &\equiv \sum_{i=0}^{\infty} d_i \left[\hat{u}_{t-i} + \rho \hat{u}_{t-i-1} + \rho^2 \hat{u}_{t-i-2} + .. \right] \end{aligned} \tag{B.2}$$

Using the definition of the $d_i's$ and factoring out identical innovations we obtain after some algebra

$$\begin{aligned} Z_t &= (1-a)\hat{z}_{t-1} + (1-a)\kappa \hat{z}_{t-2} + (1-a)\kappa^2 \hat{z}_{t-2} + .. \\ G_t &= a\hat{g}_t + (\mu a + \theta)\hat{g}_{t-1} + (\mu^2 a + \mu\theta + \theta\kappa)\hat{g}_{t-2} + (\mu^3 a + \mu^2\theta + \mu\theta\kappa + \theta\kappa^2)\hat{g}_{t-3} + .. \\ U_t &= a\hat{u}_t + (\rho a + \theta)\hat{u}_{t-1} + (\rho^2 a + \rho\theta + \theta\kappa)\hat{u}_{t-2} + (\rho^3 a + \rho^2\theta + \rho\theta\kappa + \theta\kappa^2)\hat{u}_{t-3} + .. \end{aligned} \tag{B.3}$$

where $\theta \equiv (1-a)(1-\kappa)$. Since it is a sum of innovations, the expected value of $\tilde{z}_{t|t}$ is zero. Since all the innovations are mutually and serially uncorrelated, the covariance between two adjacent forecast errors is

$$E\left[\tilde{z}_{t|t}.\tilde{z}_{t-1|t-1}\right] = E\left[Z_t.Z_{t-1}\right] + \left(\frac{\sigma_u^2}{\sigma_u^2 + \lambda^2\sigma_g^2}\right)^2 E\left[G_t.G_{t-1}\right] + \left(\frac{\lambda\sigma_g^2}{\sigma_u^2 + \lambda^2\sigma_g^2}\right)^2 E\left[U_t.U_{t-1}\right]. \tag{B.4}$$

We turn next to the calculation of the terms $E\left[Z_t.Z_{t-1}\right]$, $E\left[G_t.G_{t-1}\right]$ and $E\left[U_t.U_{t-1}\right]$. Lagging Z_t in the first equation in (B.3) by one period, multiplying by the expression for Z_t and taking the expected value of the product we obtain after some algebra

$$E\left[Z_t.Z_{t-1}\right] = (1-a)^2\kappa\left\{1+\kappa^2+\kappa^4+..\right\}^2\sigma_z^2 = \frac{(1-a)^2\kappa}{1-\kappa^2}\sigma_z^2. \tag{B.5}$$

Lagging G_t in the second equation in (B.3) by one period, multiplying the resulting expression by the expression for G_t and taking the expected value we obtain after some algebra

$$E\left[G_t.G_{t-1}\right] = \left\{\begin{array}{c} a(\mu a+\theta)+(\mu a+\theta)(\mu^2 a+\mu\theta+\theta\kappa)+ \\ (\mu^2 a+\mu\theta+\theta\kappa)(\mu^3 a+\mu^2\theta+\mu\theta\kappa+\theta\kappa^2)+.. \end{array}\right\}\sigma_g^2 \tag{B.6}$$

Since $E\left[U_t.U_{t-1}\right]$ has the same form in $\hat{u}_t$ and ρ as $E\left[G_t.G_{t-1}\right]$ has in $\hat{g}_t$ and μ it follows from (B.6) that

$$E\left[U_t.U_{t-1}\right] = \left\{\begin{array}{c} a(\rho a+\theta)+(\rho a+\theta)(\rho^2 a+\rho\theta+\theta\kappa)+ \\ (\rho^2 a+\rho\theta+\theta\kappa)(\rho^3 a+\rho^2\theta+\rho\theta\kappa+\theta\kappa^2)+.. \end{array}\right\}\sigma_u^2 \tag{B.7}$$

Equation (4.3) in the text is obtained by substituting equations (B.5) through (B.7) into equation (B.4).

C. Appendix: Proof of Remark 2

(i) The analytical expression for the derivatives of a with respect to σ_z^2/σ_g^2 and σ_z^2/σ_u^2 is rather involved and is not reported here for reason of space. We computed it using Mathematica, and verified that its value is positive for $0<\mu<1$, positive standard deviations and $\phi^2>4$ (excluding extreme cases in which at least one of the variances is zero, those conditions are always satisfied. More details on this computation are available from the authors upon request). When both ratios of variances tend to 0, T in equation (4.1) tends to zero implying, by inspection of the expression for a, that a tends to zero as well. When both ratios tend to infinity, so does T. To show that when both ratios of variances tend to infinity a tends to one, divide both the numerator and the denominator in the expression for a by T and take the limit as T goes to infinity.

(ii) Differentiating the expression for κ in equation (4.1) with respect to σ_z^2/σ_g^2

$$\frac{\partial\kappa}{\partial(\sigma_z^2/\sigma_g^2)} = \frac{\partial\kappa}{\partial\phi}\frac{\partial\phi}{\partial T}\frac{\partial T}{\partial(\sigma_z^2/\sigma_g^2)}. \tag{C.1}$$

Inspection of the expressions for κ and T shows that $\frac{\partial\kappa}{\partial\phi}<0$ and $\frac{\partial T}{\partial(\sigma_z^2/\sigma_g^2)}>0$. The derivative of ϕ with respect to T is $\frac{\partial\phi}{\partial T}=\frac{(1-\mu)^2}{(1+\mu T)^2}$ which is positive for $\mu<1$. It follows that κ is a decreasing function of σ_z^2/σ_g^2. When both variance ratios tend to zero so does T implying that ϕ tends to 2 and, therefore, that κ tends to one. The proof for σ_z^2/σ_u^2 is analogous.

D. Appendix: Numerical simulations with a Svensson (1997) type model

This section shows, by means of simulations, that results obtained in sections 3 and 4 of the text are robust to the introduction of more elaborate transmission lags and output dynamics. In particular, we consider a Svensson's (1997) type model in which the output gap and inflation are given by:

$$\begin{aligned} x_{t+1} &= \delta x_t - \varphi r_t + g_{t+1} \\ \pi_{t+1} &= \sigma \pi_t + \lambda x_t + u_{t+1} \end{aligned}$$

where x denotes the output gap, defined as $x_t \equiv y_t - z_t$ and, as in the text, the demand and cost shocks are AR(1) processes. This setup differs from the model in the text in two respects. First, it accounts for transmission lags between changes in the interest rate and economic developments. In particular, the interest rate (r_t) exerts its effects on output with a one period lag and on inflation with a two period lag. As argued by Svensson, this is consistent with VAR evidence on the effects of interest rate shocks. The second difference is that the output gap displays some degree (δ) of endogenous persistence. For computational purposes (to let the iterations on the second moments of this dynamic system converge) we remove the unit root in the potential output process assuming instead that $z_{t+1} = \gamma z_t + \hat{z}_{t+1}$ but let γ be close to one (we set $\gamma = 0.95$). For γ close to one the qualitative results concerning the policy bias and the persistence of the forecast errors due to the IP are unaffected by this modification.

Using the definitions of the output gap and the stochastic structures of the demand and cost push shock the output and inflation equations can be written as:

$$\begin{aligned} y_{t+1} &= (\gamma - \delta) z_t + \delta y_t - \varphi r_t + \mu g_t + \hat{g}_{t+1} + \hat{z}_{t+1} \\ \pi_{t+1} &= \sigma \pi_t + \lambda (y_t - z_t) + \rho u_t + \hat{u}_{t+1} \end{aligned}$$

Inspection of these equations reveals that, as in our model, observations on output contain information on demand and potential output shocks while inflation contains a signal on the cost push and potential output shocks (although the latter are lagged by one period in comparison to our case).[19]

[19] However those two signals are not quite analogous to the signals $s_{1,t}$ and $s_{2,t}$ in subsection 2.3 since, in the latter, the interest rate has been solved out in terms of variables that are known

Table D.1: Parameter values

Parameter ranges							Innovations		
μ	ρ	λ	α	γ	σ	φ	σ_z	σ_u	σ_g
(.2; .9)	(.2; .9)	.5	1.0	.95	.5	.7	.05	.15	.10

We continue to assume, as in the main text, that the information set J_t, available to the policymaker when setting r_t, includes observations up to and including the previous period's output and inflation (y_{t-1}, π_{t-1}). We assume the same objective function for monetary policy, and solve this problem using the algorithms for linear quadratic dynamic programming problems developed by Gerali and Lippi (2003). The parameters (or ranges) used in the benchmark simulations are presented in the table above.

We focus on a case where the standard deviation of potential output shocks is relatively small compared to the one of demand and cost shocks as postulated in proposition 3. As stated in that proposition (and confirmed by simulations), this is a case that gives rise to a relatively greater effects of potential output shocks. To examine the robustness of proposition 1, the simulations are conducted for two alternative cases. One, in which demand shocks are more persistent than cost shocks and and the other in which cost shocks are more persistent.

D.1. Relatively persistent demand shocks

Figure A2 presents the effects of a shock to potential output that raises actual output by one percentage point on impact.

at the beginning of period $t + 1$.

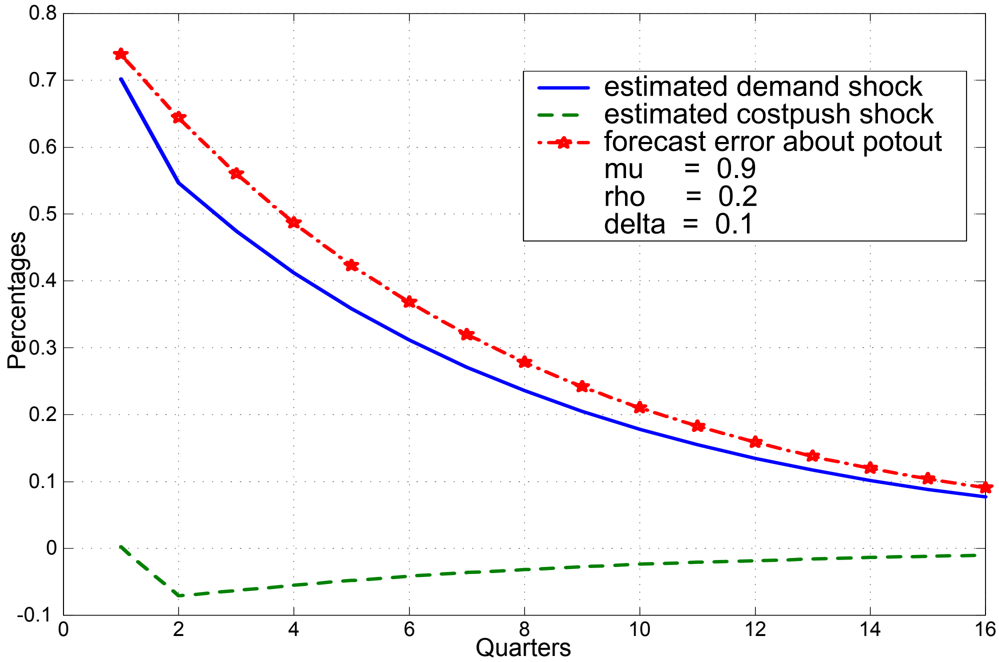

Figure A2: Perceptions following an increase in potential output.

The evolution of perceptions presented in the figure is based on the assumption that $\mu = 0.9$ and $\rho = 0.2$ implying that the signal extraction problem is dominated by the persistence of demand shocks. The potential output shocks hits the economy, i.e. output y_o, at $t = 0$ (horizontal axis). Given the assumed lags in the information structure, the policymaker observes this output at $t = 1$; the relatively small variance of potential output shocks leads the policymaker to attribute a large part of the observed output increase to a demand shock, as occurred in the model of the main text. The solid line in the figure shows this fact.

In addition, the low inflation observed in the subsequent period leads the policymaker to infer the presence of a negative cost push shock (dashed line).[20] The persistence of errors in forecasting potential output is illustrated by the dashed-starred line; it takes about 4 quarters to learn half of the potential output shock. Figure A3 thus confirms proposition 1 in the context of the extended model for relatively persistent demand shocks.

[20]Notice how the transmission lags assumed in this model induce the forecast errors about the cost push shock to "appear" with a one period lag in comparison to the model in the text.

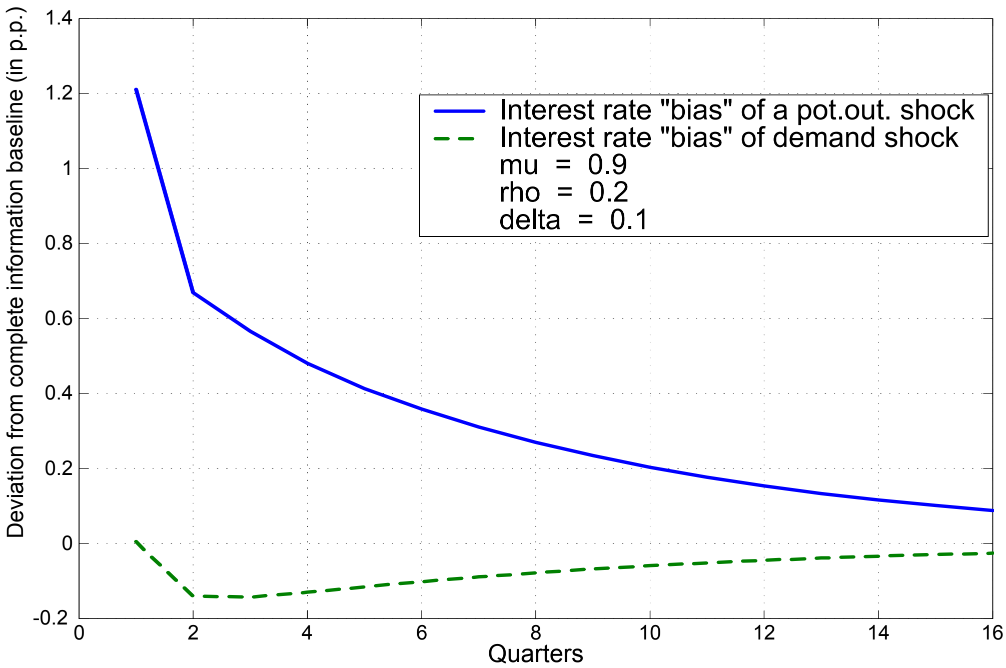

Figure A3: Policy bias with relatively persistent demand shocks.

When demand shocks are more persistent than cost shocks, policy is dominated by "demand" considerations. Following a potential output shock, which would require no response of the interest rate under complete information, policy rates are raised to offset the perceived positive output gap. As can be seen from figure A3 this results in a persistent excessively tight monetary policy. The figure also shows that, given the relatively low variance of potential output innovations, a one percent shock to output originating from a potential output shock induces a much larger deviation of the interest rate (from its full information benchmark) than an idential output innovation originating from a demand shock of (dashed line).

D.2. Relatively persistent cost shocks

The next two figures consider the opposite case, in which cost push shocks are more persistent than demand shocks ($\mu = 0.2$ and $\rho = 0.9$). Figure A4 shows that a less persistent demand shock induces the policymaker to attribute less of the observed output increase to demand shocks which reduces the forecast errors in potential output in comparison to the previous case (compare the dash-starred lines across Figures A2 and A4).

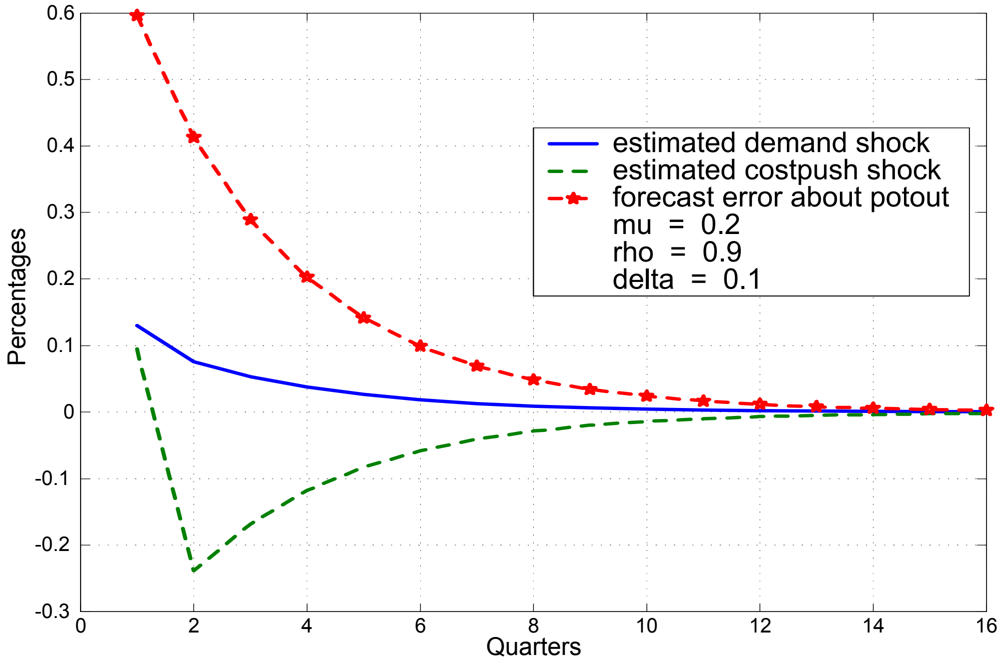

Figure A4: Perceptions following an increase in potential output.

In addition, since the cost shock is relatively persistent, the size of the perceived cost shock is larger than in the first case (compare the broken lines in Figures A2 and A4). As discussed in proposition 1, in such a case policy is dominated by "cost-push" considerations so that, as soon as a cost push shock is perceived, (i.e. from the second quarter onwards) policy turns out to be excessively loose rather than excessively tight as predicted by the second part of proposition 1. This behavior is illustrated in Figure A5.

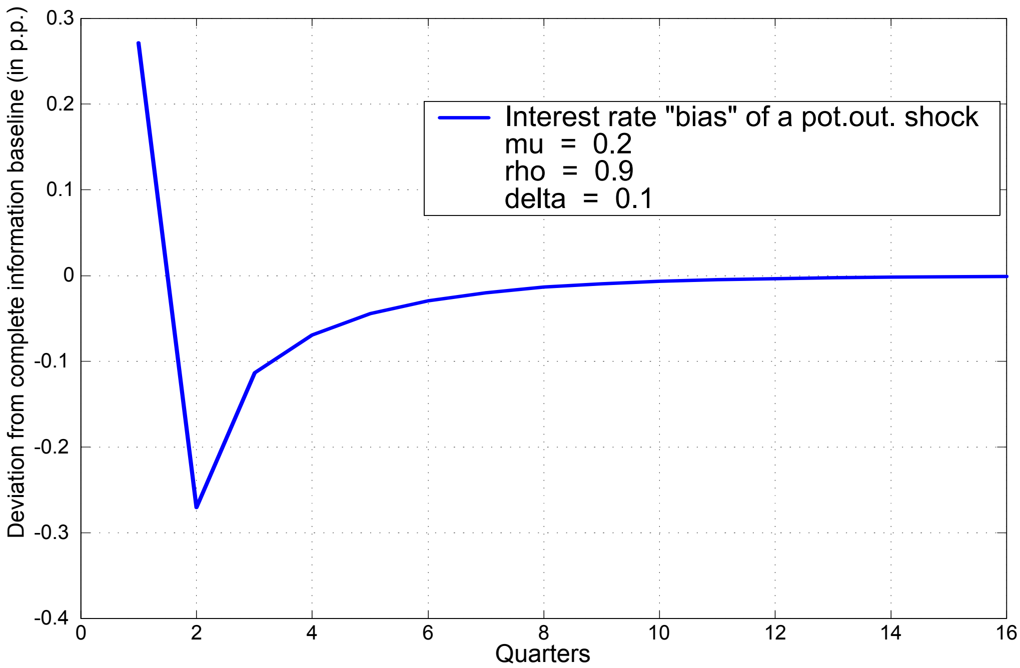

Figure A5: Policy bias with relatively persistent cost shocks.

D.3. Conclusions

The main lesson from the simulations with the more elaborate structure of lags and endogenous persistence is that propositions 1, 2 and 3 appear to be qualitatively robust to the introduction of such extensions.

References

[1] Blanchard O. and Quah D., 1989, "The Dynamic Effects of Aggregate Demand and Supply Disturbances", **American Economic Review**, 79: 655-673.

[2] Brunner K., Cukierman A. and Meltzer A., 1980, "Stagflation Persistent Unemployment and the Permanence of Economic Shocks", **Journal of Monetary Economics**, 6: 467-492.

[3] Brunner K. and Meltzer A., 1993, "Money and the Economy: Issues in Monetary Analysis. The Raffaele Mattioli Lectures", Cambridge: Cambridge University Press.

[4] Clarida R., Galí J. and Gertler M., 2000, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory", **Quarterly Journal of Economics**, 113:147-180.

[5] Cukierman A., 1984, *Inflation, Stagflation, Relative Prices and Imperfect Information,* Cambridge University Press, Cambridge, London, New York.

[6] Cukierman A. and Meltzer A., 1982, "What Do Tests of Market Efficiency in the Presence of the Permanent - Transitory Confusion Show?" Unpublished Manuscript. Available at: http://www.tau.ac.il/~alexcuk/pdf/Cukierman_and_Meltzer_1982.pdf

[7] Cukierman A., 1998, "The Economics of Central Banking", in Wolf Holger (ed.), *Contemporary Economic Issues Macroeconomics and Finance*, V.5, The MacMillan Press, (IEA conference volume 125), pp. 37-82.

[8] A. Cukierman, G. P. Miller and B. Neyapti, 2002, "Central Bank Reform, Liberalization and Inflation in Transition Economies - An International Perspective", **Journal of Monetary Economics**, 49: 237-264.

[9] Economic Report of the President, 1979, United States Government Printing Office, Washington.

[10] Ehrmann, M. and F. Smets, 2003. "Uncertain potential output: implications for monetary policy", **Journal of Economic Dynamics and Control**, 27: 1611-1638.

[11] Gerali A. and Lippi F., 2002, "Optimal Control and Filtering in Linear Forward-Looking Economies: A Toolkit", mimeo, Bank of Italy.

[12] Hansen L. and Sargent T., 1997, *Recursive Linear Models of Dynamic Economies,* Manuscript, Available at: http://www.stanford.edu/~sargent/hansen.html

[13] Hamilton, J.D, 1994, *Time Series Analysis*, Princeton University Press, Princeton.

[14] Kuttner K.N., 1992, "Monetary Policy with Uncertain Estimates of Potential Output", Federal Reserve Bank of Chicago Economic Perspectives, 16: 2-15.

[15] Kuttner K.N., 1994, "Estimating Potential Output as a Latent Variable", **Journal of Business & Economic Statistics**, 12: 361-368.

[16] Lansing K., 2000, "Learning about a Shift in Trend-Output: Implications for Monetary Policy and Inflation", Federal Reserve Bank of San Francisco, Working Papers in Applied Economic Theory and Econometrics, No. 16.

[17] Orphanides A., 2001, "Monetary Policy Rules Based on Real-Time Data", **American Economic Review**, 91: 964-985.

[18] Orphanides A., 2003a, "The Quest for Prosperity Without Inflation", **Journal of Monetary Economics**, 50: 633-663.

[19] Orphanides A., 2003b, "Historical monetary policy analysis and the Taylor rule", **Journal of Monetary Economics**, 50: 983-1022.

[20] Rogoff K., 1985, "The Optimal Degree of Commitment to a Monetary Target", **Quarterly Journal of Economics**, 100:1169-1190.

[21] Rudebush, G., D., 2002. "Assessing nominal income rules for monetary policy with model and data uncertainty", **Economic Journal**, 112: 1-31.

[22] Siklos P., 2002, *The Changing Face of Central Banking*, Cambridge University Press, Cambridge UK and NY.

[23] Svensson L.E.O., 1997. "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets". **European Economic Review**, 41: 1111-1146.

[24] Swanson E.T., 2004, "Signal Extraction and Non-Certainty-Equivalence in Optimal Monetary Policy Rules", **Macroeconomic Dynamics**, 8: 27-50.

[25] Taylor J.B., 1998, "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank", **Journal of Monetary Economics**, 43: 655-679.

[26] Walsh C., 1995, “Optimal Contracts for Independent Central Bankers”, **American Economic Review**, 85: 150-167.