Endogenous Monetary Policy with Unobserved Potential Output*

Alex Cukierman† and Francesco Lippi‡

December 31, 2002

Abstract

This paper characterizes endogenous monetary policy when policymakers are uncertain about the extent to which movements in output and inflation are due to changes in potential output or to cyclical demand and cost shocks. We refer to this informational limitation as the "information problem" (IP). Main results of the paper are: 1. Policy is likely to be excessively loose (restrictive) for some time when there is a large decrease (increase) in potential output in comparison to a full information benchmark. This provides a partial but unified explanation for the inflation of the seventies and the price stability of the nineties. 2. Errors in forecasting potential output and the output gap are generally serially correlated. 3. A quantitative assessment, based on an empirical model of the US economy developed by Rudebusch and Svensson (1999) indicates that, during and following periods of large changes in potential output, the IP significantly affects the dynamics of inflation and output. 4. The increase in the Fed's conservativeness between the seventies and the nineties, and a more realistic appreciation of the uncertainties surrounding potential output in the second period, imply that the IP problem had a stronger impact in the seventies than in the nineties.

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 Board of Governors of the Fed, the Bank for International Settlements and at the Central European University. We benefitted from extremely useful discussions with Athanasios Orphanides and from the comments of Simon Gilchrist and David Small.

†Amnon Ben-Nathan Professor of Economics, Tel - Aviv University and Center, Tilburg University. E-mail: alexcuk@post.tau.ac.il

‡Research Department, Bank of Italy. E-mail: lippi.francesco@insedia.interbusiness.it

1. Introduction

A stabilizing role for monetary policy crucially hinges on some notion of "potential output", a non-observable economic variable representing the desirable level at which actual output should be. The conduct of monetary policy requires, therefore, 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 relevance of uncertainty about potential output for real-time policymaking. He examined the difficulties inherent in real time estimation of potential output and suggested that because of signal extraction errors that arise under imperfect information, situations requiring policy actions may not be recognizable until later on.

This policy implication is central for Orphanides (2000a,b, 2001), who reports evidence of a significant (real time) overestimation of potential output during the oil shocks of the seventies. Based on this finding Orphanides argues that, by leading to a monetary policy stance which turned out to be excessively loose with the benefit of hindsight, this overestimation 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, again with the benefit of hindsight, that the greater-than-expected increases in potential output could have allowed a less restrictive monetary policy stance than the stance initially implied by real time estimates of inflation and of the output gap.

The hypothesized relevance of imperfect information may shed interesting 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 the evaluation of the extent to which (retrospective) policy errors were avoidable in real time. This paper makes a step in this direction by proposing such a benchmark model and by analyzing its properties.

The paper shows that, given the structure of information, some policy decisions, which are judged ex-post to be mistakes, may be unavoidable even if the central bank utilizes the best forecasting procedures in “real time”. Such 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. Those claims are established in an environment where the central bank cannot perfectly disentangle (not even ex-post) between changes in inflation and output that are due to changes in potential output from those that are due to higher frequency changes in demand and costs. We label this inevitable confusion between demand and cost shocks, on one hand, and shocks to potential output, on the other, as an “information problem” or IP in brief.[1]

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. The analytical framework of this paper provides an “optimizing” analytical foundation for this mechanism and identifies the conditions under which it operates.[2] Interestingly, a large permanent decrease in potential output does not lead to an excessively loose and persistent policy stance under all circumstances. Whether it does or not depends on the relative persistence of demand and of cost shocks, on the degree of conservativeness of the central bank and on the relative size of the variance of innovations to

[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.

[2] Related work in which potential output is specified as a Hodrick-Prescott filter appears in Lansing (2000). Two differences between our paper and that of Lansing are that in our paper the forecast of potential output is derived from the stochastic structure of the economy, and the policy rule is derived from the loss function of policymakers. By contrast, Lansing postulates both of those concepts exogenously.

The paper by Swanson (2000) 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. By contrast we focus on the implications of such a framework for optimal interest rate policy and for the associated retrospective ”policy errors”.

potential output.

The results above are developed within a simple macroeconomic model which underlies the conception of many central banks about the transmission process of monetary policy.[3] The paper identifies conditions under which the presence of 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 relatively to this benchmark in periods of large expansions in potential output. The 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, as was the case in the seventies, there is a large decrease in potential output, policymakers interpret part of this reduction as a negative output gap and loosen monetary policy too much in comparison to a no IP benchmark. Thus, in periods of large decreases in potential output, inflation accelerates partly because of the relatively expansionary monetary policy stance. Conversely, when − as might have been the case in the US during the nineties − a "new economy" raises the level of potential output, inflation goes down partly because policy makers interpret part of the increase in output as a positive output gap and thus policy is tighter than under perfect information.

A main novel result of the paper is that, even when the information available to policymakers in real time is used efficiently and monetary policy chosen optimally, the forecast errors in real time estimates of potential output and of the output gap are serially correlated retrospectively. 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 serial correlation is due to innovations to potential output. Interestingly, retrospective evidence about forecast errors in potential output during the seventies and the eighties are consistent with these implications (Orphanides, 2000a). As a consequence of the serial correlation in those errors monetary policy also appears to retrospectively be systematically biased in one direction.

In summary the paper provides a unified framework for understanding some of the reasons for the inflation of the seventies, as well as for the remarkable price

[3]It is a compact formulation of the economic structure that appears in Svensson (1997).

stability of the nineties. 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 wether imperfect information is quantitatively important for the determination of monetary policy and inflation. In section 5 we make a first step in this direction by using the economic structure estimated by Rudebusch and Svensson (1999) to obtain a preliminary estimate of the quantitative impact of the IP problem under optimal monetary policy and optimal filtering. Finally, the paper argues that it is likely that the IP problem was less important during the nineties than during the seventies for two reasons. First the Fed was more conservative in the latter period. Second both the economic profession and the Fed had a more realistic evaluation of the uncertainties surrounding potential output in the latter period.

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 rates of interest, 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 illustrates the main prediction of the theoretical model by means of a numerical analysis. It then uses an empirical model of the US economy, proposed by Rudebusch and Svensson (1999), to provide a preliminary quantitative assessment of the effects of imperfect information. Section 6 discusses reasons supporting the view that retrospective policy errors were smaller during the nineties than during the seventies. This is followed by concluding remarks.

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 similar to those in Svensson (1997) and Rudebusch and Svensson (1999). Despite the model's simplicity, it is likely to capture key elements of the views of major central banks about the transmission mechanism of monetary policy.[4] Its main advantage is that it illustrates analytically some basic consequences of imperfect information in a relatively simple manner. The main qualitative effects described in the analytical sections of the paper also appear in richer models, featuring transmission lags (as in the Rudebusch Svensson model considered in Section 5.1) and forward-looking variables (as in Ehrmann and Smets (2001) and Gerali and Lippi, [2002]).

2.1. The economy

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

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

$$\pi_t = \lambda(y_t - z_t) + u_t. \tag{2.2}$$

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. This framework postulates that potential output z is a fundamental long run determinant of actual output. But, in addition, 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.

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 like technology and the accumulation of physical and human capital.[5] The permanence of shocks to potential output is modeled by assuming that z_t is a random walk.[6] More specifically we postulate the following stochastic

[4] Mishkin's (1999) comment on Rudebusch and Svensson (1999) elaborates the sense in which such a, relatively simple, model is useful for monetary policy makers.

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

[6] Nothing in our results would change if we had added a deterministic trend growth to the

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. The first order condition for this problem implies

$$x_{t|t-1} = -\frac{\lambda}{\alpha} \pi_{t|t-1}. \tag{2.7}$$

Here $x_{t|t-1}$ and $\pi_{t|t-1}$ are the expected values of inflation and of the output gap 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, observations on actual inflation and output up to and including period $t-1$. A full specification of J_{t-1} appears below. Since period's t values of inflation and of the output gap are not known with certainty at the beginning of period t, those variables (which are indirectly controlled by policy) appear in equation (2.7) in expected terms.

The equilibrium outcomes for the interest rate, output and inflation obey

potential output process.

(Appendix A):

$$
\begin{aligned}
r_t &= \frac{1}{\varphi}\left[g_{t|t-1} + \frac{\lambda}{\alpha+\lambda^2}u_{t|t-1}\right] && (2.8)\\
y_t &= z_t + (g_t - g_{t|t-1}) - \frac{\lambda}{\alpha+\lambda^2}u_{t|t-1} && (2.9)\\
\pi_t &= \frac{\alpha}{\alpha+\lambda^2}u_t + \lambda\left(g_t - g_{t|t-1}\right) + \frac{\lambda^2}{\alpha+\lambda^2}\left(u_t - u_{t|t-1}\right). && (2.10)
\end{aligned}
$$

2.3. The structure of information and optimal policy

The interest rate rule in (2.8) 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

$$J_t = \{\Omega, y_{t-i}, \pi_{t-i,} \mid i = 0, 1, 2, ...\} \tag{2.11}$$

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 (2.9) and (2.10)):

$$
\begin{aligned}
s_{1,t} &\equiv y_t + g_{t|t-1} + \frac{\lambda}{\alpha+\lambda^2}u_{t|t-1} = z_t + g_t && (2.12)\\
s_{2,t} &\equiv \pi_t + \lambda g_{t|t-1} + \frac{\lambda^2}{\alpha+\lambda^2}u_{t|t-1} = \lambda g_t + u_t && (2.13)
\end{aligned}
$$

where variables to the left of the equality sign are observed separately while those to the right are not.[7] 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 (2.12) and (2.13), once the optimal estimate of potential output, $z_{t|t}$, is known.[8] 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:

$$\tilde{u}_{t|t} \equiv u_t - u_{t|t} \tag{2.14}$$

$$\tilde{g}_{t|t} \equiv g_t - g_{t|t} \tag{2.15}$$

$$\tilde{z}_{t|t} \equiv z_t - z_{t|t} \tag{2.16}$$

Using (2.12) and (2.13) the following useful relationship between these errors can be derived :

$$\lambda \tilde{z}_{t|t} = -\lambda \tilde{g}_{t|t} = \tilde{u}_{t|t}. \tag{2.17}$$

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.[9] This is summarized in the following remark.

[7] 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 $t+1$.

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

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

$$g_{t|t} = g_t - \tilde{z}_{t|t} \tag{2.18}$$

$$u_{t|t} = u_t + \lambda \tilde{z}_{t|t}. \tag{2.19}$$

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 reference to equations (2.12) and (2.13). 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 about cyclical conditions (cost-push and demand shocks). The purpose of this subsection is to answer the following question: How do such noisy perceptions about the cycle affect monetary policy, inflation and the output gap? We do this 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 policy makers possess, at the beginning of period $t+1$, 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 (2.8), (2.18), (2.19) and (2.17), the *deviation* of the optimal interest rate in the presence of the IP from its optimal value under full information (i.e. $r^*_{t+1} = \frac{1}{\varphi}\left[\mu g_t + \frac{\lambda}{\alpha+\lambda^2}\rho u_t\right]$) can be written as

$$\Delta r_{t+1} \equiv r_{t+1} - r^*_{t+1} = -\frac{1}{\varphi}\left[\mu \tilde{g}_{t|t} + \frac{\lambda\rho}{\alpha+\lambda^2}\tilde{u}_{t|t}\right] \tag{3.2}$$

$$= \frac{\left(\mu - \rho\frac{\lambda^2}{\alpha+\lambda^2}\right)}{\varphi}\tilde{z}_{t|t}. \tag{3.3}$$

It follows immediately from (3.2) that if demand shocks are sufficiently persistent in comparison to cost shocks (i.e. $\mu > \frac{\rho\lambda^2}{\alpha+\lambda^2}$) 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 ($\mu > \frac{\rho\lambda^2}{\alpha+\lambda^2}$) 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 ($\mu < \frac{\rho\lambda^2}{\alpha+\lambda^2}$)

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 $t+1$.[10] 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).

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 analyze the deviations of outcomes obtained in the presence of the IP from those that arise in its absence. Using (2.9) and (2.10) it is immediate to relate these

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

deviations to the interest rate deviations studied above. This yields:

$$\Delta x_{t+1} \equiv x_{t+1} - x_{t+1}^* = -\varphi \Delta r_{t+1} \tag{3.4}$$

$$\Delta \pi_{t+1} \equiv \pi_{t+1} - \pi_{t+1}^* = -\varphi \lambda \Delta r_{t+1} \tag{3.5}$$

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 (2000, 2001) empirical results according to which, during the seventies, US monetary policy was overly expansionary due to an overestimation 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. 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.

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 convey the intuition of the basic mechanisms at work we focus in the text on the particular (but simpler) case in which 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.2. It is shown there that the main qualitative properties of the optimal predictor when shocks are equally persistent carry over to the more general case.

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

This subsection describes the optimal predictor of potential output in the case in which 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 B.1):[11]

$$
\begin{aligned}
z_{t|t} &= aS_t + (1-a)(1-\kappa)\sum_{i=0}^{\infty}\kappa^i S_{t-1-i} && (4.1)\\
\text{where } &: \\
\kappa &\equiv \tfrac{2}{\phi+\sqrt{\phi^2-4}} \in (0,1) \qquad \phi \equiv \tfrac{2+T(1+\mu^2)}{1+\mu T} \geq 2;\ T \equiv \left(\tfrac{\sigma_z^2}{\sigma_g^2} + \tfrac{\lambda^2\sigma_z^2}{\sigma_u^2}\right) \\
a &\equiv \tfrac{[(1-\mu)+(1-\kappa)+T(1-\mu\kappa)]T}{[T(1-\mu-\mu\kappa)+(1-\mu-\kappa)](1+T)+(T+\mu)(1+\mu T)} \in (0,1) && (4.2)\\
S_{t-i} &\equiv s_{1,t-i} - \tfrac{\lambda\sigma_g^2}{\sigma_u^2+\lambda^2\sigma_g^2} s_{2,t-i} = z_{t-i} + \tfrac{\sigma_u^2\cdot g_{t-i}-\lambda\sigma_g^2\cdot u_{t-i}}{\sigma_u^2+\lambda^2\sigma_g^2} && (4.3)
\end{aligned}
$$

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 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 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 note that although the true value of potential output is contained only in the signals $s_{1,t-i}$, the optimal predictor assigns positive weights **also** 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 $s_{1,t-i}$.

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

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 up 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 a large number of 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 one of the cyclical components 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 types of 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, contrary to forecast errors of variables which 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 manner that errors of forecast of a variable that is revealed one period after the formation of that forecast, is normally used to update future forecasts.[12]

[12] 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

As a matter of fact it can be shown that forecast errors of potential output and of 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 (2.17), that the error in forecasting the output gap is equal to minus the error of forecast in 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 C that the covariance between two adjacent forecast errors is given by

$$
\begin{aligned}
E\left[\tilde{z}_{t|t}.\tilde{z}_{t-1|t-1}\right] = {} & \frac{(1-a)^2\kappa}{1-\kappa^2}\sigma_z^2 + \\
& \left(\tfrac{\sigma_u^2}{\sigma_u^2+\lambda^2\sigma_g^2}\right)^2 \left\{ \begin{array}{c} a(\mu a+\theta)+(\mu a+\theta)(\mu^2 a+\mu\theta+\theta\kappa)+ \\ (\mu^2 a+\mu\theta+\theta\kappa)(\mu^3 a+\mu^2\theta+\mu\theta\kappa+\theta\kappa^2)+.. \end{array} \right\} \sigma_g^2 \\
& +\left(\tfrac{\lambda\sigma_g^2}{\sigma_u^2+\lambda^2\sigma_g^2}\right)^2 \left\{ \begin{array}{c} a(\rho a+\theta)+(\rho a+\theta)(\rho^2 a+\rho\theta+\theta\kappa)+ \\ (\rho^2 a+\rho\theta+\theta\kappa)(\rho^3 a+\rho^2\theta+\rho\theta\kappa+\theta\kappa^2)+.. \end{array} \right\} \sigma_u^2
\end{aligned} \tag{4.4}
$$

where

$$\theta \equiv (1-a)(1-\kappa). \tag{4.5}$$

Since, except for the extreme case in which $a = 0$ and $\kappa = 1$ all terms on the right hand side of equation (4.4) are positive, errors in forecasting potential output exhibit a positive serial correlation. 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 (2000a). 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

that is being forecasted is not revealed with certainty even after the fact, forecast errors are serially correlated in general.

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

Lemma 1. *(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.*

(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 D. An immediate implication of the Lemma 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 (4.5) is not far from zero. But inspection of equation (4.4) reveals that when a and θ are not far from zero the coefficients of σ_g^2 and of σ_u^2 in equation (4.4) 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 go up and the coefficient of σ_z^2 goes down.

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 relatively 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 the lemma) 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 gradual, 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 several periods as a persistent change in the output gap.

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

Proposition 2 implies that the serial correlation is always present in the population. But it will be particularly in evidence following the realization of a large change in potential output. The reason is that, in finite samples, the magnitude

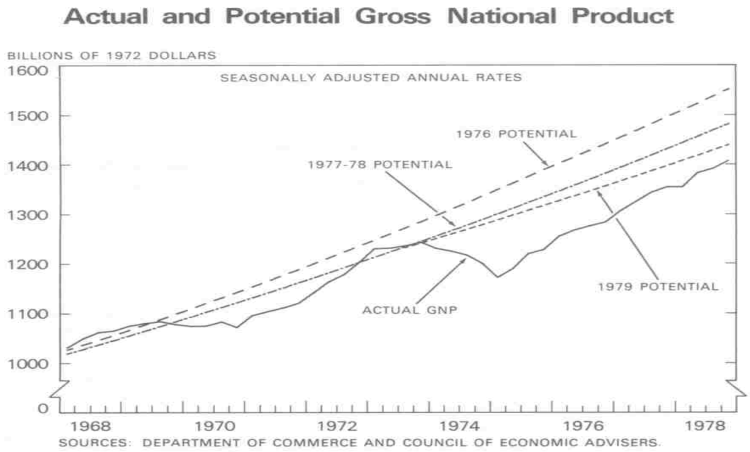

Figure 4.1: Reprint from "Economic Report of the President (1979)".

of the serial correlation is directly related to the size of the shock to potential output.[13] 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 gradually but optimally.

This point of view fits surprisingly well the persistent downward revisions of estimates of potential output in the US during the latter part of the seventies. Enlightening documentation on this persistent process of backward downward revisions of perceived potential output appears in the 1979 Economic Report of the President (pp. 72-76.). In particular, chart 7 (which we report) vividly illustrates the magnitude and persistence of this process. 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.*

(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

[13]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.

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 (2000b, 2001). 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 partly interpreted as a positive output gap and triggered a policy response that was judged retrospectively to be overly restrictive.

5. A Quantitative Illustration

As a practical illustration of the effects described above, we present an impulse response analysis of the effects of a potential output shock under imperfect information.[14] We parametrize our model economy using the parameters reported in Table 5.1 below. This example illustrates the impulse responses of the main variables in the system to a one percent shock to potential output. Figure 5.1 illustrates how the signal extraction problem faced by the policy maker under imperfect information creates a confusion about the sources of business cycle fluctuations.

The upper box in the figure displays the true pattern followed by the (unit root) potential output (z_t) after the shock. The estimated potential output pattern $z_{t|t}$ (computed with the Kalman filter) is traced out in the second box of the figure. In line with the theory of the previous section, the learning process is gradual and the forecast errors exihibit positive serial correlation. The three remaining boxes in the Figure illustrate how missperceptions about potential output translate into

[14]The numerical implementation of this exercise relies on the algorithms discussed in Gerali and Lippi (2002).

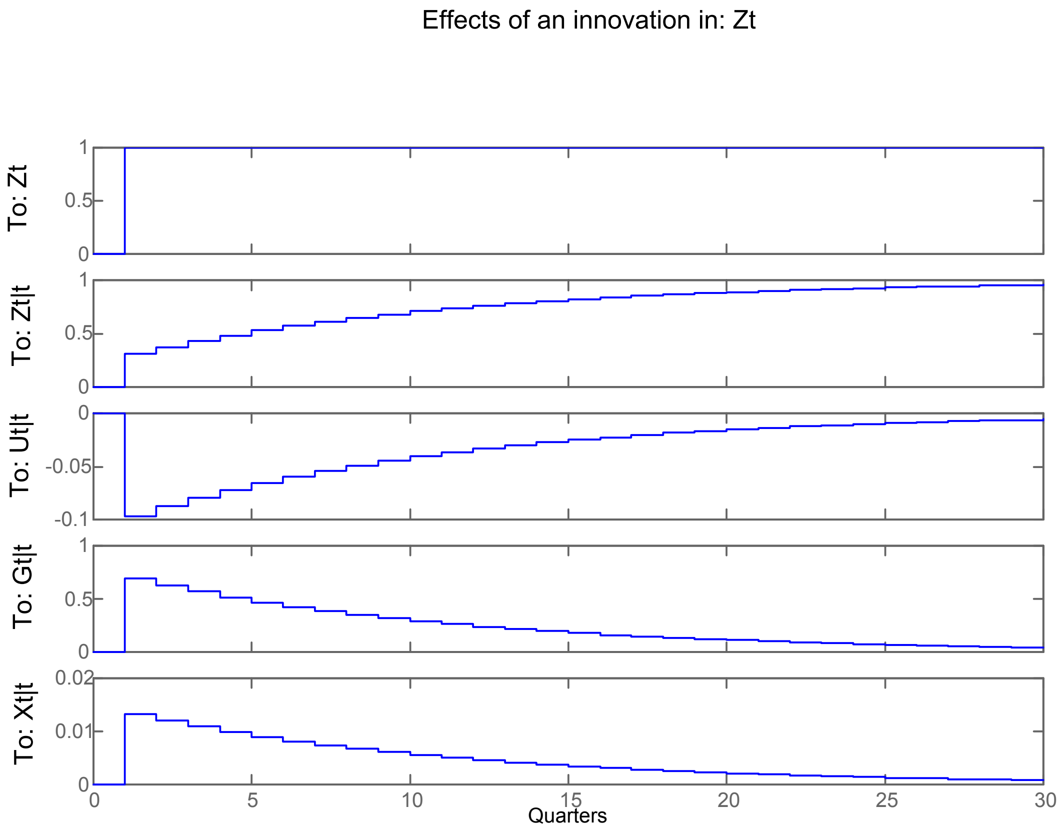

Figure 5.1: Perceived state of the economy in response to a PO shock

Table 5.1: Baseline parameter values

Parameters					
β	α	λ	ρ	μ	φ
.99	1	.14	.7	.7	1

Innovations (std)		
σ_z	σ_u	σ_g
.5	1.5	1.5

missperceptions about the cost-push shock ($u_{t|t}$), the demand shock ($g_{t|t}$) and the output gap ($x_{t|t}$), all of which are identically zero in this experiment (these relationships obey equation (2.17)). It is evident that an underestimated potential output level implies an overestimated demand shock (to "explain" the currently high observed output level) and an underestimated cost shock (consistent with the relatively low realized inflation). Quantitatively, of the true 1 per cent increase in potential output only 0.3 percent are perceived initially, while the remainder (about 0.7 percent) is erroneously interpreted as a demand shock, and therefore, as an output gap.

The macroeconomic consequences of these missperceptions are depicted in Figure 5.2. The parameters chosen are such that the inequality $\mu > \frac{\rho\lambda^2}{\alpha+\lambda^2}$ is satisfied, implying that monetary policy is driven mainly by "demand shock" considerations (see Proposition 1). Recall that, under the complete information benchmark, there should be no policy response following this shock, i.e. the optimal interest rate path should be equal to the optimal interest rate in the absence of cyclical shocks, which is identically zero. The figure shows how, under imperfect information, a positive innovation to potential output causes the interest rate to rise above its optimal value under perfect information. This causes the true output gap (x_t) to become negative (although the policy maker perceives a positive output gap, see figure 5.1) and inflation to be lower than under the full information benchmark. This pattern fits a situation like the "nineties", during which high output growth is associated with low inflation.

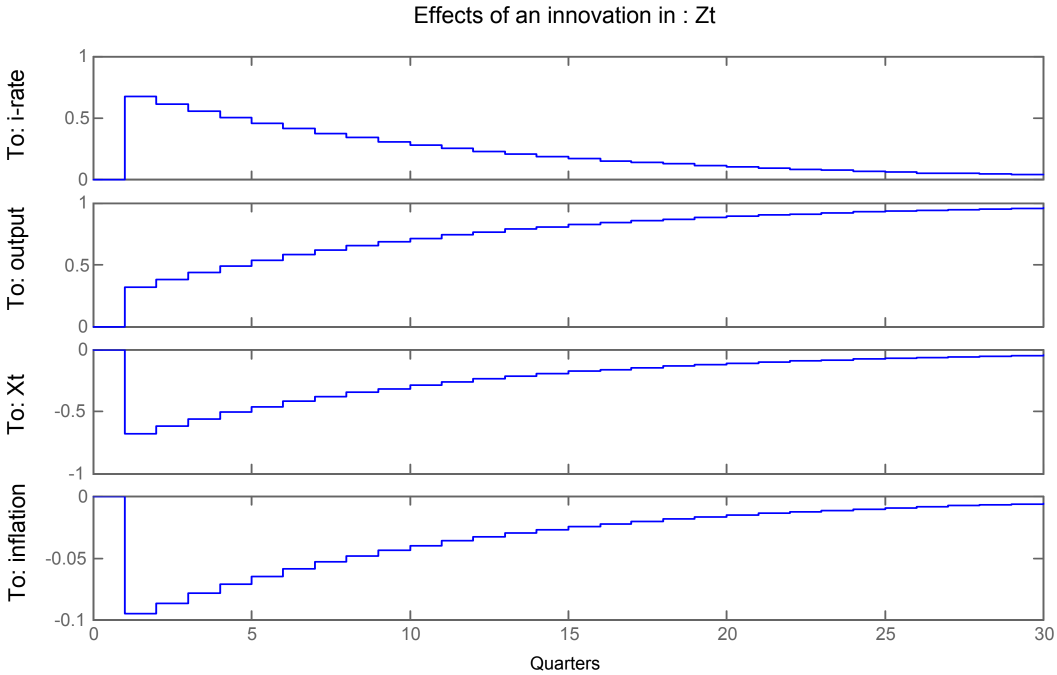

Figure 5.2: Macro effects of a PO shock with imperfect info

5.1. A quantitative assessment based on the Rudebusch and Svensson (1999) model.

For a more realistic assessment of the quantitative effects of imperfect information we repeat the exercise developed above using a model of the US economy proposed and estimated by Rudebusch and Svensson (1999). The model consists of the following two autoregressive equations for inflation (π_t) and the output gap (x_t):

$$\pi_t = \sum_{j=1}^{4} \beta_j \pi_{t-j} + \lambda x_{t-1} + \hat{u}_t \tag{5.1}$$

$$x_t = \sum_{j=1}^{2} \alpha_j x_{t-j} - \phi r_{t-1} + \hat{g}_t \tag{5.2}$$

where $\hat{u}_t$ and $\hat{g}_t$ are white noise processes. The parameters $\alpha_j, \beta_j, \lambda, \phi$ and the standard deviations of inflation and of output gap innovations are estimated by OLS. Rudebusch and Svensson (1999, p.208) argue that, despite its simplicity, this model provides a description of the US economy which, from the perspective of monetary policy, conforms with the received wisdom encapsulated in the MPS model, "which was used regularly in the Federal Reserve's forecasting process over 25 years."[15]

Our objective here is to analyze the consequences of a non-observable potential output level. Recall that we define the output gap as the percentage difference between actual and potential output ($x_t \equiv y_t - z_t$). We postulate that equation (5.2) originates from the output equation $y_t = \sum_{j=1}^{2} \alpha_j y_{t-j} - \phi r_{t-1} + \hat{g}_t + \hat{z}_t$ and from the potential output equation $z_t = \sum_{j=1}^{2} \alpha_j z_{t-j} + \hat{z}_t$. This formulation implies that, as in the main body of the paper, potential output shocks ($\hat{z}_t$) do not affect the output gap (as they shift actual and potential output in the same way). Therefore they should not be stabilized by the policy maker. An information problem exists, however, as potential output and cyclical shocks to demand and inflation cannot be observed separately .

We follow Rudebusch and Svensson by assuming that the policy maker aims at minimizing an intertemporal loss function $\Lambda_t = E[\sum_{\tau=0}^{\infty} \beta^{\tau} L_{t+\tau} \mid I_t]$,where the

[15]The estimated coefficients in Rudebusch and Svensson (1999) are: $\alpha_1 = 1.16$, $\alpha_2 = -.25$, $\beta_1 = .70$, $\beta_2 = -.10$, $\beta_3 = .28$, $\beta_4 = .12$, $\lambda = .14$, $\phi = .10$ (the units are annual percent values for inflation and percentages for the output gap variable).

period loss function is given by:

$$L_t \equiv \left[\alpha\,(x_t)^2 + (\pi_t)^2 + \nu(r_t - r_{t-1})^2\right]$$

and adopt their basic parametrization in setting $\alpha = 1$ and $\nu = .5$. The last term reflects the well documented tendency of central banks to adjust interest rates in small steps. Imperfect information enters the policy problem through our assumption that period's t information set, J_{t-1}, includes only observations on actual output and inflation up to and including period $t-1$ and no direct observations on either past or current levels of potential output. As discussed in Section 4.3, the signal to noise ratios σ_z^2/σ_g^2 and σ_z^2/σ_u^2 are key in determining the outcomes of the filtering problem. The experiments below utilize the values estimated by Rudebusch and Svensson (1999), respectively $\sigma_u = 1.009$ and $\sigma_g = .819$ (annual percent values for inflation and percentage points for the output gap). As to the innovations in potential output, we experimented with values ranging from "small" ($\sigma_z = 0.1$) to "large" ($\sigma_z = 1.0$). As implied by Proposition 3, following the realization of an isolated shock to potential output, the forecast errors in potential output are larger and more persistent in the former case (small σ_z). This case, which is discussed below, provides the most favourable setting for potential output shocks to create large and persistent "policy errors". The reason is that, in this case, the signal about potential output in observable data is already so small that the forecast of potential output is largely insensitive to new information and is, therefore, nearly a constant.[16]

Figure 5.3 reports the deviations between the paths of four main macro variables under imperfect and under perfect information following a 1 percentage point reduction in potential output. The upper box in the figure shows that the policy maker's forecast error of the output gap is initially very large (almost none of the shock is predicted) and highly persistent (it takes about four years to return near zero). The interest rate is lower than under full information, as almost all of the output reduction is perceived as a cyclical shock. As a consequence, both output and inflation are above their full information counterpart (lower box of the figure).

[16] As a consequence further reductions in σ_z do not lead to noticeable changes in the effects of imperfect information.

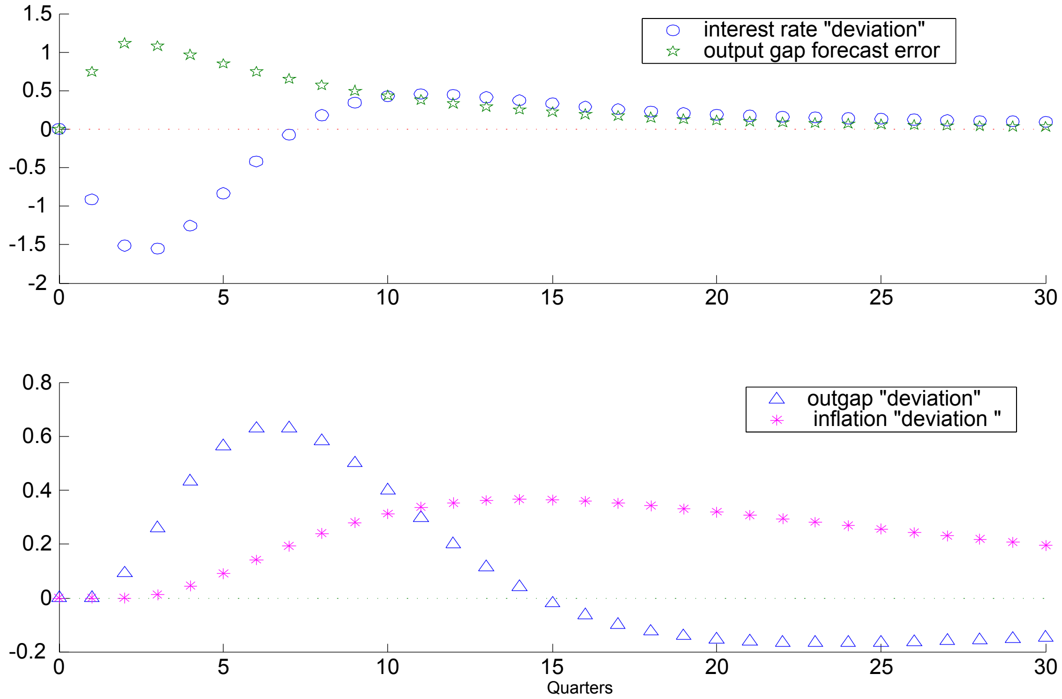

Figure 5.3: Effects of imperfect information following a negative potential output shock

A back-of-the envelope calculation can be used to gauge the economic significance of the magnitudes predicted by our model. The revisions in the estimates of potential output for the seventies reported in Figure 4.1 suggest that forecast errors in the output gap are in the range of 4 to 7 percent of output (for the year 1976). Somewhat larger magnitudes are suggested by Orphanides's (2000, Figure 3) measures of the forecast errors in the output gap for the 1970s. If we choose a benchmark value of about 5 percent for the error in the output gap, we have to scale all the effects in Figure 5.3 up by a factor of 5. This implies that the interest rate under incomplete information is more than five percentage points below its full information counterpart during the year following the shock. This calculation also indicates that inflation and the output gap record maximum deviations of about 2 and 3.5 percentages from their full information benchmarks respectively. While those numbers are economically significant, indicating that imperfect information might contribute to explain the higher than average inflation recorded in the mid seventies, they admittedly only go part of the way, leaving a significant part of that inflationary burst to be explained by other factors. Three potential candidates are the direct inflationary impact of the oil shocks, inefficient use of real time information and/or inefficient implementation of monetary policy.

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

Taken literally the previous analysis implies that, other things 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 have changed between the seventies and the nineties.

First the relative emphasis of policy on stabilization of inflation versus output stabilization shifted towards stabilization of inflation. In terms of our model this means that the parameter α has decreased between the seventies and the nineties implying, via equation (2.8), that the response of the interest rate to cost shocks in the nineties is stronger than in the seventies. Arguments and evidence presented in Taylor (1998), Clarida, Galí and Gertler (2000) and Siklos (2002, pp. 61-64) supports this view. Second, it is likely that during the seventies policymakers had an overly optimistic view of their ability to forecast potential output and the natural level of employment. The view that potential output is rather difficult to predict became accepted mainly during the nineties as illustrated, *inter alia*, by the work of Staiger, Stock and Watson (1997a, 1997b). 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.

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

Proposition 1 implies that, provided $\mu > \rho\lambda^2/(\alpha+\lambda^2)$, 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 has been 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 $\mu - \rho\lambda^2/(\alpha + \lambda^2)$. Since during the nineties policy has been relatively more conservative, $\alpha_{70s} > \alpha_{90s}$, which implies that

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

This leads to the following proposition

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 is similar during the seventies and during the nineties, policy errors should be 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) as well as 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.

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

We embed the notion that during the seventies policymakers were overoptimistic about potential output uncertainty into the analysis by postulating that during this period the perceived variance, σ_{zp}^2, of the innovation to potential output was lower than the true variance, σ_z^2, but that during the nineties the perceived variance adjusted upward and became equal to the true variance. Other than that 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, 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, that a is

[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_{zp}^2, rather than with the actual variance, σ_z^2.

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 **also** because of a swifter and more accurate recognition of changes in potential output.

7. Concluding remarks

This paper provides a unified explanation for part of the inflation of the seventies, as well as for part of the remarkable price stability of the nineties. This is done by showing that, even if monetary policy is optimal and forecasts of potential output 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 the extent to which policy was excessively loose in the seventies is larger than the extent to which it was excessively tight during the nineties. This conclusion is based on two presumptions and associated consequences.

The first presumption is that the Fed was relatively more conservative in Rogoff (1985) sense in the nineties than in the seventies. The second presumption is that, due to a relatively more realistic evaluation of uncertainties surrounding potential output, the Fed learned more quickly and more accurately about changes in potential output during the nineties than during the seventies. The first effect is due to the fact that, given 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 mechanism is due to the fact that, since it learned about changes in potential output more quickly and accurately during the nineties, the Fed's policy was nearer to the full information optimal policy also because of more appropriate forecasting procedures.

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 those errors are not too important, nor are they 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 substantial and to be serially correlated. 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 from this statement 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, it does not necessarily follow from the above conclusion that policy and forecasting procedures during the seventies were as efficient as possible at the time. The point, however, is that it is not possible to conclude just from the ex-post identification of policy errors that such errors were avoidable in real time. The real challenge facing policymakers and economists is to distinguish between avoidable (in real time) and unavoidable policy errors. We believe that a model like the one proposed here, where policy is consistent with the economic structure and information is processed efficiently, can pave the way towards facing such a challenge.

The second conclusion is that, with the exception of extreme cases, the fact that, during periods following 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.

A. Appendix: Model Solution

Condition (2.7) implies the interest rate rule:

$$r_t = \frac{1}{\varphi}\left[g_{t|t-1} + \frac{\lambda}{\alpha}\pi_{t|t-1}\right] \tag{A.1}$$

which yields the following output and inflation outcomes:

$$y_t = z_t + (g_t - g_{t|t-1}) - \frac{\lambda}{\alpha}\pi_{t|t-1} \tag{A.2}$$

$$\pi_t = \lambda\left[\left(g_t - g_{t|t-1}\right) - \frac{\lambda}{\alpha}\pi_{t|t-1}\right] + u_t \tag{A.3}$$

Note that (A.3) contains an expected inflation term which, by the rational expectations hypothesis, is:

$$\pi_{t|t-1} = \frac{\alpha}{\alpha+\lambda^2}u_{t|t-1} \tag{A.4}$$

B. 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 $s_{3,t-i} \equiv s_{1,t-i} - s_{2,t-i}$. Let us write the linear predictor for z_t conditional on J_t as:

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

and the last line follows immediately from (2.12) and (2.13). 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 $\min_{a_i,b_i} Q$,where:

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

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

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

and

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

Note that the two 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 (B.3) by one step, multiplying the resulting expression by μ and subtracting it from (B.3) yields:

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

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

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

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

$$\begin{aligned} o &= -\sigma_z^2\left[(a_{i+1}+b_{i+1})-\mu(a_{i+2}+b_{i+2})\right]+ \\ &\quad +\sigma_g^2\left[(1-\mu)^2(\mu^i a_0+...+a_i)-(2-\mu)a_{i+1}+a_{i+2}\right] \end{aligned} \tag{B.7}$$

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

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

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

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

where both (B.8) and (B.9) hold for $i=1,2,3,...$. 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 the simpler case in which $\mu=\rho$ and then present the general solution.

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

When $\mu=\rho$ the difference equations (B.8) and (B.9) can be uncoupled. It is immediate to see that in such case the a_i and b_i are related by the linear relationship

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

where the equality for $i=0$ is established from the first order conditions for a_0 and b_0 (not reported). Substituting the expression for the generic b_i into (B.8) yields

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

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

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

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: $\kappa^2 - \phi\kappa + 1$ (from B.11). The values of a_0 and of a_1 remain to be determined. Using the first order conditions for a_0 and a_1 (where the latter is obtained from (B.3) for $i = 1$) the following linear relation is established (after some algebraic transformations of identical nature to those used to establish (B.8)):

$$a_1 \equiv \frac{(1-\mu)(1+T)a_0 - \frac{\sigma_z^2}{\sigma_g^2}(1-\mu)}{(1+\mu T)}. \tag{B.13}$$

A second linear relation between a_0 and a_1 is established after analogous algebraic transformations are applied to equation (B.5) for $i = 1$. This yields

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

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

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

Some algebra reveals that $a_0 + \frac{a_1}{1-\kappa} = \left(1 + \frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)$, which suggests the convenient reformulation of the filter used in the main text which is based on the modified signal $S_{t-i} \equiv S_{t-i}^{'}\left(1 + \frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)^{-1}$. Under this formulation, rewrite (B.15) using $a \equiv a_0\left(1 + \frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)^{-1}$ and $a_1^{'} \equiv a_1\left(1 + \frac{\lambda^2\sigma_g^2}{\sigma_u^2}\right)^{-1}$. Since $a + a_1^{'}/(1-\kappa) = 1$, this implies $a_1^{'} = (1-\kappa)(1-a)$ used in (4.1) in the main text.

B.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 (B.8) and (B.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

$$x_{t+1} = Ax_t + Cw_{t+1} \tag{B.16}$$

where

$$x_{t+1} \equiv \begin{bmatrix} z_{t+1} \\ g_{t+1} \\ u_{t+1} \end{bmatrix}, A \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \rho \end{bmatrix}, C \equiv \begin{bmatrix} \sigma_z & 0 & 0 \\ 0 & \sigma_g & 0 \\ 0 & 0 & \sigma_u \end{bmatrix}, \tag{B.17}$$

and where w_{t+1} is a vector of iid innovation with unit variance. The system in equation (B.16) is the Kalman filter's state equation. Rewriting equations (2.12) and (2.13) in matrix form we obtain

$$y_t = Gx_t + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{B.18}$$

where

$$y_t \equiv \begin{bmatrix} s_{1,t} \\ s_{2,t} \end{bmatrix}, G \equiv \begin{bmatrix} 1 & 1 & 0 \\ 0 & \lambda & 1 \end{bmatrix}. \tag{B.19}$$

Equation (B.18) 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 (B.16) and (B.18) 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

$$x_{t|t} = x_{t|t-1} + K\left[y_t - Gx_{t|t-1}\right] \tag{B.20}$$

where

$$K \equiv \Sigma G' \left[G\Sigma G'\right]^{-1} \tag{B.21}$$

[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.

and

$$\Sigma = A\Sigma A' + CC' - A\Sigma G' \left[G\Sigma G'\right]^{-1} G\Sigma A'. \tag{B.22}$$

Equation (B.22) implicitly determines the unknown matrix, Σ, and given Σ, equation (B.21) determines K. Equation (B.20) can be rewritten as

$$x_{t|t} = \left[I - KG\right] x_{t|t-1} + Ky_t. \tag{B.23}$$

Lagging (B.23) by one period and using $x_{t+1|t} = Ax_{t|t}$, repeated substitution of the resulting expression into (B.23) yields

$$x_{t|t} = \sum_{j=0}^{\infty} D^j K y_{t-j} \tag{B.24}$$

where

$$D \equiv \left[I - KG\right] A, \;\; D^0 \equiv I \tag{B.25}$$

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 (B.24), the optimal predictor of potential output can be written as

$$z_{t|t} = \sum_{j=0}^{\infty} k_{11}^j S_{t-j} \tag{B.26}$$

where

$$S_{t-j} \equiv s_{1,t-j} + \omega^j \cdot s_{2,t-j}, \;\; j = 0, 1, ...\infty. \tag{B.27}$$

$$\omega^j \equiv \frac{k_{12}^j}{k_{11}^j} \tag{B.28}$$

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 to .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 B1 plots the coefficients k_{11}^j for the first six lags

Table B.1: Baseline parameter values

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

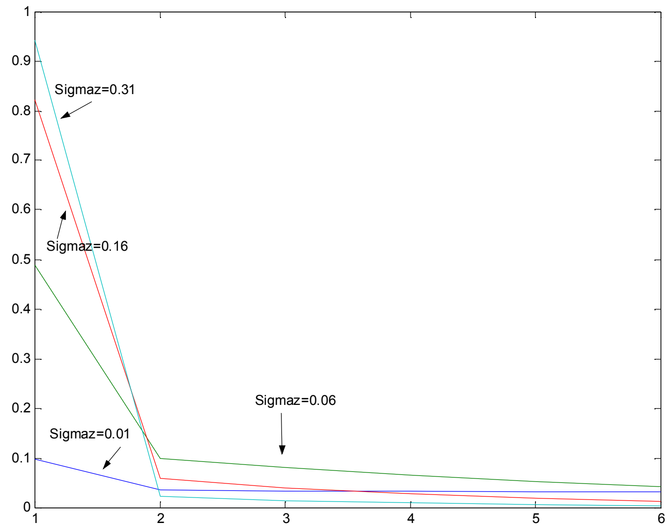

Figure B.1: Weights k_{11}^j on Observables for $j = 0, 1, .., 5$.

($j = 0, 1, .., 5$) computed from the optimal filter for four different values of σ_z (ranging from relatively small, $\sigma_z = 0.01$, to relatively large, $\sigma_z = 0.31$).

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 gets old. 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 larger (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, becomes smaller as σ_z increases.

C. Appendix: Investigation of the serial correlation properties of errors of forecast of 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 (2.16) and regrouping terms so as to express this error in terms of infinite sums of the innovations in z, g and u we obtain

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

where

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

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

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

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 therefore

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

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 (C.3) by one period, multiplying by the expression for Z_t and taking the expected value of the product we obtain after some algebra

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

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

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

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 (C.6) that

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

Equation (4.4) in the text is obtained by substituting equations (C.5) through (C.7) into equation (C.4).

D. Appendix: Proof of Lemma 1

(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 one or more 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

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

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.

References

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

[2] Brunner Karl, Alex Cukierman and Allan Meltzer, 1980, "Stagflation Persistent Unemployment and the Permanence of Economic Shocks", **Journal of Monetary Economics**, 6, October, 467-492.

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

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

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

[6] Cukierman Alex and Allan Meltzer, 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] Economic Report of the President, 1979, United States Government Printing Office, Washington.

[8] Ehrmann, Michael and Frank Smets, 2001 "Uncertain potential output: implications for monetary policy", ECB working paper No.59.

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

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

[11] Kuttner, Kenneth N., 1992, "Monetary policy with uncertain estimates of potential output", *Federal Reserve Bank of Chicago Economic Perspectives*, 16, January-February 2-15.

[12] Kuttner, Kenneth N., 1994, "Estimating potential output as a latent variable", *Journal of Business & Economic Statistics*, Vol. 12, No.3, pp.361-368.

[13] Lansing, K. 2000, “Learning about a shift in trend-output: implications for monetary policy and inflation”, mimeo, Federal Reserve Board of Governors.

[14] Meyer L. H. 2000, Remarks by Governor Laurence Meyer before the Joint Conference of the Fed of San Francisco and the Stanford Institute for Economic Policy Research, March 3.

[15] Mishkin, F., 1999, .“Policy Rules for Inflation Targeting: Comment”, in *Monetary Policy Rules*, J. Taylor (Ed.), The university of Chicago Press, Chicago and London.

[16] Orphanides, A., 2000a, ”Activist Stabilization Policy and Inflation: The Taylor Rule in the 1970s”, manuscript, Board of Governors of the Federal Reserve System.

[17] Orphanides, A., 2000b, “The quest for prosperity without inflation”, manuscript, Board of Governors of the Federal Reserve System.

[18] Orphanides, A., 2001, “Monetary Policy Rules based on Real-Time Data”, **American Economic Review**, Vol.91(4): 964-985.

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

[20] Rudebusch G.D. and L.E.O. Svensson, 1999, “Policy Rules for Inflation Targeting”, in *Monetary Policy Rules*, J. Taylor (Ed.), The university of Chicago Press, Chicago and London.

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

[22] Staiger D., Stock J. H. and Watson M. W., 1997a, “The NAIRU, Unemployment and Monetary Policy”, **Journal of Economic Perspectives**, 11, (Summer):33-50

[23] Staiger D., Stock J. H. and Watson M. W., 1997b, “How Precise Are Estimates of the Natural Rate of Unemployment?” in Romer C. D. and Romer D. H. (eds.) **Reducing Inflation : Motivation and Strategy**, University of Chicago Press, Chicago.

[24] Svensson, Lars O.E., 1997. “Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets”. **European Economic Review**, 41:1111-46.

[25] Svensson, Lars O.E., 1999. “Inflation Targeting as a Monetary Policy Rule”. **Journal of Monetary Economics**. Vol. 43:607-654.

[26] Svensson, Lars O.E. and Michael Woodford, 1999. “Implementing Optimal Policy through Inflation-Forecast Targeting”. Mimeo, Princeton University, NJ.

[27] Svensson, Lars O.E. and Michael Woodford, 2000. “Indicator variables for optimal policy”. Mimeo, Princeton University, NJ.

[28] Swanson Eric T., 2000, ”On Signal Extraction and Non-Certainty-Equivalence in Optimal Monetary Policy Rules”, Board of Governors of the Federal Reserve System.

[29] 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.

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