

TRANSPARENCY AND CREDIBILITY:  
MONETARY POLICY WITH  
UNOBSERVABLE GOALS

Jon Faust  
Lars E. O. Svensson

Working Paper **6452**

NBER WORKING PAPER SERIES

TRANSPARENCY AND CREDIBILITY:  
MONETARY POLICY WITH  
UNOBSERVABLE GOALS

Jon Faust  
Lars E. O. Svensson

Working Paper 6452  
<http://www.nber.org/papers/w6452>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
March 1998

The authors thank participants in seminars at the Board of Governors, the Institute for International Economic Studies, Reserve Bank of Australia, University of Canterbury and Victoria University of Wellington, and especially David Bowman, Martin Flodén, Dale Henderson, Andy Levin, Bennett McCallum, Stefan Palmqvist, Torsten Persson and Michael Woodford for helpful comments. They also thank Christina Lönnblad for secretarial and editorial assistance. Part of the paper was written when Lars Svensson visited the Reserve Bank of New Zealand and Victoria University of Wellington; he thanks these institutions for their hospitality. Remaining errors are the authors' own; the views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, the Reserve Bank of New Zealand, other members of their staffs or the National Bureau of Economic Research.

© 1998 by Jon Faust and Lars E. O. Svensson. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Transparency and Credibility: Monetary  
Policy with Unobservable Goals  
Jon Faust and Lars E. O. Svensson  
NBER Working Paper No. 6452  
March 1998  
JEL Nos. E52, E58

### ABSTRACT

We define and study transparency, credibility, and reputation in a model where the central bank's characteristics are unobservable to the private sector and are inferred from the policy outcome. A low-credibility bank optimally conducts a *more* inflationary policy than a high-credibility bank, in the sense that it induces higher inflation, but a *less* expansionary policy in the sense that it induces lower inflation and employment than expected. Increased transparency makes the bank's reputation and credibility more sensitive to its actions. This has a moderating influence on the bank's policy. Full transparency of the central bank's intentions is generally socially beneficial, but frequently not in the interest of the bank. Somewhat paradoxically, direct observability of idiosyncratic central bank goals removes the moderating incentive on the bank and leads to the worst equilibrium.

Jon Faust  
Board of Governors  
of the Federal Reserve System  
20th and Constitution Avenues, NW  
Washington, DC 20551  
faustj@frb.gov

Lars E. O. Svensson  
Institute for International Economic Studies  
Stockholm University  
S-10691 Stockholm  
SWEDEN  
and NBER  
lars.svensson@iies.su.se

# Transparency and Credibility: Monetary Policy with Unobservable Goals\*

Jon Faust<sup>†</sup> and Lars E.O. Svensson<sup>‡</sup>

First draft: June 1997

This version: December 1997

## Abstract

We define and study transparency, credibility, and reputation in a model where the central bank's characteristics are unobservable to the private sector and are inferred from the policy outcome. A low-credibility bank optimally conducts a *more* inflationary policy than a high-credibility bank, in the sense that it induces higher inflation, but a *less* expansionary policy in the sense that it induces lower inflation and employment than expected. Increased transparency makes the bank's reputation and credibility more sensitive to its actions. This has a moderating influence on the bank's policy. Full transparency of the central bank's intentions is generally socially beneficial, but frequently not in the interest of the bank. Somewhat paradoxically, direct observability of idiosyncratic central bank goals removes the moderating incentive on the bank and leads to the worst equilibrium.

## 1 Introduction

In December 1989, as U.S. inflation was cresting 5 percent for the first time in 6 years, the Federal Open Market Committee (FOMC) held discussions regarding whether the Fed should more firmly pursue "price stability."<sup>1</sup> FOMC members generally agreed that price stability was

---

\* The authors thank participants in seminars at the Board of Governors, the Institute for International Economic Studies, Reserve Bank of Australia, University of Canterbury and Victoria University of Wellington, and especially David Bowman, Martin Flodén, Dale Henderson, Andy Levin, Bennett McCallum, Stefan Palmqvist, Torsten Persson and Michael Woodford, for helpful comments. They also thank Christina Lönnblad for secretarial and editorial assistance. Part of the paper was written when Lars Svensson visited the Reserve Bank of New Zealand and Victoria University of Wellington; he thanks these institutions for their hospitality. Remaining errors are the authors' own; the views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System, the Reserve Bank of New Zealand, or other members of their staffs.

<sup>†</sup> Board of Governors of the Federal Reserve System, [faustj@frb.gov](mailto:faustj@frb.gov), <http://www.patriot.net/users/faustj>.

<sup>‡</sup> Institute for International Economic Studies, Stockholm University, [Lars.Svensson@iies.su.se](mailto:Lars.Svensson@iies.su.se), <http://www.iies.su.se>; CEPR and NBER.

<sup>1</sup> We put price stability in quotation marks, since when central bankers refer to price stability they may mean low or zero inflation, which implicitly or explicitly allows base drift of the price level. The price level in this case has a unit root and, outside central banking circles, would probably not be thought of as "stable."

their inflation goal, with FOMC Vice Chairman Corrigan referring to their [[11], p. 45] “collective zeal” on this point. When President Forrestal (Atlanta Fed) questioned public support for raising unemployment to lower inflation from just under 5 percent, Dr. Prell (director of the Fed’s Division of Research and Statistics) immediately responded that [p. 14–15] “...if the public thinks that the FOMC is thinking this way, then that means there is no credibility to the disinflationary commitment.... So we’re in that credibility bind....” Several members offered views like that of FOMC vice chairman Corrigan [p. 30–31]: “...I don’t think it’s prudent for this institution... to bet the ranch on that [credibility] because if we’re wrong we’ve got a heck of a problem on our hands...” President Stern (Minneapolis Fed) stated [p. 21]: “...I personally would start with the weak credibility case.... [I]f you start with something as pessimistic as that I think you have a difficult challenge in a rigorous way to justify [the pursuit of price stability].” The FOMC chose not to pursue its zealously-held goal at that time.

This discussion provides a dramatic example of how thoroughly the concepts from literature on commitment and discretion—beginning with Kydland and Prescott [18] and Barro and Gordon [3]—have been integrated into positive and normative discussions of monetary policy. Both policymakers and academics now discuss policy in the language of game-theoretic political economy. This paradigm rose to favor in industrialized countries because it offered an account of why governments in the 1960s and onwards chose widely unpopular and historically high rates of inflation. The large literature that has followed Kydland and Prescott and Barro and Gordon has spawned a sort of *folk wisdom* regarding credibility and transparency. This wisdom involves claims like<sup>2</sup>

1. A central bank with low credibility should, everything else equal, conduct a more restrictive policy than a high-credibility bank,
2. A central bank with low credibility has less flexibility to respond optimally to shocks, and
3. Greater transparency improves credibility and thereby policy outcomes.

For instance, following the third wisdom, several countries have attempted to raise the credibility and transparency of low-inflation policy by announcing explicit inflation targets and by issuing regular *Inflation Reports*.<sup>3</sup>

Since the early-1980s, many countries have begun to attain extended records of low-to-moderate inflation. While the concepts of credibility, reputation, and transparency continue to

---

<sup>2</sup> See for instance Federal Reserve Bank of Kansas City [10] for similar statements. Bernanke and Mishkin [4] is a standard reference to statement (2). King [17] argues for (3).

<sup>3</sup> See Leiderman and Svensson [19], Haldane [15] and Federal Reserve Bank of Kansas City [10].

play a prominent role in policy analysis, it is not clear that the literature provides clear guidance as to the relevance of the concepts under moderate inflation. After six years of low inflation, should a central bank still expect to have low credibility? Should the reaction to low credibility be a looser policy than under high credibility (as the FOMC chose) or tighter policy as suggested by the folk wisdom? How does increased transparency improve performance, and is increased transparency both in the interest of society and of the central bank?

The goal of this paper is formally to assess the importance of dynamic credibility and transparency under persistently low inflation (say, below five percent). The most easily attained goal of such work is to provide a laboratory in which to assess the validity of the folk wisdom claims. Our results indicate, for instance, that a low-credibility bank conducts a *more* expansionary policy than a high credibility bank in the sense that it induces higher inflation, but a *less* expansionary policy in the sense that it induces lower-than-expected inflation, and, hence, lower employment. As for transparency, the results are more subtle. Increased transparency of the central bank's intentions is generally socially beneficial. Such transparency makes the bank's reputation and credibility more sensitive to its actions. This has a moderating influence on the bank's policy. Full transparency of the bank's intentions is generally socially optimal, but for interesting cases, this transparency is not in the interest of the central bank. Surprisingly, what we call extreme transparency, when the central bank's idiosyncratic goals are directly observable, removes any moderating incentive on the bank and results in the worst possible outcome both for the bank and for the public. We believe these results capture important generalities, but the model still is missing too many elements of reality to be taken directly to the data at this stage.

Section 2 specifies the main building blocks and the basic features of our model. Section 3 specifies our baseline regime, when the central bank's goals and intentions are unobservable to the private sector, as well as to two alternative regimes. Section 4 summarizes how credibility affects optimal policy; section 5 discusses transparency. Section 6 summarizes and concludes. The appendices contain technical details.

## 2 Building blocks

Episodes like that of the FOMC in 1989 suggest two features that must be captured in a model of dynamic credibility in moderate inflation countries. First, the public continuously attempts to deduce the central bank's preferences from policy outcomes. Modelling this requires that

the central bank’s relative preferences for employment and inflation evolve in ways not fully observed by the public. In reality, central bank preferences represent an aggregation of the heterogeneous preferences in society as manifested through the selection process of central bankers, the policymaker’s interactions within the rules for decision-making, and their reaction to external pressures.<sup>4</sup> At moderate rates of inflation, the marginal trade-off between inflation and unemployment for key groups in society is unclear, making the central bank’s process of preference aggregation extremely opaque. We assume that this can be approximated as a central bank with goals that are to some extent both random and unobservable from the public’s perspective.

Second, the central bank’s policy actions do not flawlessly reveal its intentions. This feature works hand-in-hand with the first feature in guaranteeing that the public is not certain about the central bank’s taste for inflation and employment. We believe that there is sufficient noise in the economy so that, e.g., when the inflation rate briefly rose above 5 percent in 1990 in the U.S. there was understandable dissent in the public over whether the Fed had a greater taste for inflation than previously thought.

The model of Cukierman and Meltzer [7] (CM) is an excellent starting point for work exhibiting these two features. In CM, the central bank has unobserved goals that evolve continuously over time. More precisely, a parameter in the central bank’s loss function—which can be interpreted as the marginal gain from increased output—is dynamic, stochastic, and private information of the central bank. Further, noise in the economy keeps the public from flawlessly deducing the central bank’s intentions from economic observables.

An alternative is to start with a model in which the central banker’s type is drawn—perhaps from a discrete set—and remains fixed for a fixed term. When the central banker’s type is fixed for all time, there is no scope at all for studying the role of variation of preferences. When the term is finite (as in Backus and Driffill [1]), the important dynamics of credibility and reputation arise from the closeness to the end of term, at which point the central banker spends any remaining reputation and reveals her true type. We know of little empirical support for the view that end-of-term dynamics are crucial.<sup>5</sup>

While the CM model has great potential, and we believe deserves far more attention than it has received, it has drawbacks that we attempt to address. First, the CM loss function is questionable, since it can be interpreted as being linear in output: the central bank would accept

---

<sup>4</sup> See Faust [8] for an examination of the importance of rules of selection and bargaining for monetary policy in the U.S.

<sup>5</sup> Several factors may limit the importance of such dynamics, for example, overlapping terms and the importance of reputation after leaving the post.

arbitrary increases in employment variance for tiny reductions in inflation. As we show, this choice of loss function implies that central banks behave the same way regardless of whether credibility is low or high. This prediction of the indifference of central banks regarding credibility overturns claims (1) and (2) above and is strongly counterfactual. Second, the effects of transparency in CM are inextricably linked with control-error variance—unavoidable error in implementing policy decisions—so that improving transparency also means improving monetary control. We would like to capture aspects of certain real world efforts to improve transparency, which often involve measures, such as issuing inflation reports, that are intended to increase transparency but do not directly alter the degree of monetary control.

## 2.1 The model

The model differs formally from CM's only in the period loss function and in the specification of the control error. The model has two agents, the private sector (also called the public) and the central bank. The private sector's behavior is summarized by two relations. First, employment is generated by a standard Phillips curve,

$$l_t = (\pi_t - \pi_{t|t-1}) + \varepsilon_t, \quad (2.1)$$

where  $l_t$  is (log) employment in period  $t$ , and  $\pi_t$  is the inflation rate in period  $t$  (the change in the log price level between period  $t - 1$  and  $t$ ) and  $\varepsilon_t$  is an employment shock (a supply shock). The average rate of employment,  $E[l_t]$ , is normalized to equal zero. The second relation defines  $\pi_{t|t-1}$ , the private-sector expectation of inflation. Private-sector expectations are rational, in that their expectation of inflation is the mathematical expectation in the model, given the private sector information set. In order to keep track of the asymmetric information, our conventions are that the expectations operator with respect to central bank information is  $E$ , and that with respect to private-sector information is  $E^p$ . Subscripts like  $_{t|t-1}$  always indicate the private-sector conditional expectation of a variable in period  $t$ , seen from  $t - 1$ ; thus, given the rational expectations assumption,  $\pi_{t|t-1} = E_{t-1}^p \pi_t$ .

The central bank has imperfect control over inflation,

$$\pi_t = i_t + \eta_t, \quad (2.2)$$

where  $i_t$  is the central bank's *intention* for inflation, and  $\eta_t$  is a mean-zero control error. Note that  $i_t$  is not the central bank's instrument, which is usually easily observable. This equation



sweeps under the rug all issues of how the central bank's intention is (imperfectly) implemented through the manipulation of its instruments. In our baseline case, the central bank's intention is not observed by the public, and (2.2) captures the realistic feature that outcomes observed by the public do not fully reveal central bank intentions.<sup>6</sup>

The central bank's loss function at the end of period  $t - 1$  is

$$\mathbb{E}_{t-1} \sum_{j=t}^{\infty} \beta^{j-t} L_j, \quad (2.3)$$

where  $\beta$  ( $0 < \beta < 1$ ) is a discount factor, and where the period  $t$  loss function is

$$L_t \equiv \frac{1}{2} \left[ \pi_t^2 + (l_t - l_t^*)^2 \right]. \quad (2.4)$$

The central bank's total employment target,  $l_t^*$ , fulfills

$$l_t^* = l^* + z_t, \quad (2.5)$$

$$z_t = \rho z_{t-1} + \theta_t, \quad (2.6)$$

where  $l^* \geq 0$  is the long-run employment target,  $z_t$  is a time-varying preference parameter that we call the employment target,  $0 \leq \rho < 1$ , and  $\theta_t$  is a shock to the target. We will generally take  $z_t$  as being unobservable to the public.

These preferences can be interpreted as representing a central bank with an explicit zero inflation target, and an implicit, unobservable, and time-varying employment target. We interpret the stochastic portion of the loss function as arising from shifts in the way the central banking structure aggregates heterogeneous societal preferences.<sup>7</sup> Given heterogeneous individual preferences in the private sector, it is clear that (2.4) does not represent an unambiguous measure of social loss; indeed, we are presuming that there is no such measure. Because it is common to assume that central banks have preferences that are in some way *unrepresentative* of the public,<sup>8</sup> we will make some welfare comparisons using what we stipulate to be a more *representative* social loss function. In particular, we view the  $z_t$  component of (2.4) as representing idiosyncratic central-bank preferences (that may, for instance, reflect a temporarily increased influence or pressure from some special-interest group) and we say that the representative social loss function has the same form as (2.3), but with the period loss given by

$$L_t^p \equiv \frac{1}{2} \left[ \pi_t^2 + (l_t - l^*)^2 \right], \quad (2.7)$$

---

<sup>6</sup> At the expense of added mathematical complication, we could add an instrument that the central bank manipulates and that is observed by the public but whose value does not directly reveal intentions.

<sup>7</sup> It might seem natural to have  $l^*$  fixed but to have the relative weights on the inflation and employment terms vary stochastically. Under this formulation, however, the solution to the problem is not a linear decision rule.

<sup>8</sup> Cf., for instance, Rogoff [27], Walsh [30] and [31], Persson and Tabellini [25] and Svensson [29].

where  $z_t$  does not enter. Thus, the private sector appoints a central banker with goals that coincide with the private sector's on average, but that wander a bit with the vagaries of the central bankers' preferences.<sup>9</sup>

While we examine several regimes, in all regimes the central bank has full information about its preferences and, at the end of period  $t$ , it has full information about all period  $t$  shocks. The time line in each period goes as follows. At the end of period  $t - 1$ , the public forms its expectations of period  $t$  variables. The central bank observes those expectations. At the beginning of period  $t$ , the central bank observes its employment target,  $z_t$ , and the supply shock,  $\varepsilon_t$ , and chooses its intention,  $i_t$ . Next, the control error,  $\eta_t$ , is realized, giving  $\pi_t$  and the public observes  $\varepsilon_t$ , giving  $l_t$ . Then the cycle begins again. All of the shocks in the model are jointly normal, mutually uncorrelated, and have zero mean and fixed variance. The variance of any particular shock  $x$  is denoted  $\sigma_x^2$ .

## 2.2 Reputation, credibility, and transparency

One of the benefits of this framework is that it allows fairly natural implementations of the key notions of credibility, reputation, and transparency. These concepts all relate to the dynamics of central-bank preferences. When  $z_t$  is high, extra employment is especially valuable to the bank, raising the marginal benefit to positive inflation surprises. If  $z_t$  were fully observed in period  $t - 1$ , then when the public observed a high  $z_t$  its inflation expectation would rise to the point at which the marginal employment benefit to the central bank from slightly higher than expected inflation would just be offset by the inflation cost. The bank would then implement the expected inflation rate.

When  $z_t$  is not observable, the public's inflation expectation depends on its period  $t - 1$  assessment of the central bank's period  $t$  employment target,  $z_{t|t-1}$ . Thus,  $z_{t|t-1}$  summarizes the *reputation* of the central bank among private sector agents.<sup>10</sup>

We are also interested in the *credibility* of the zero (expected) inflation policies that are optimal in our model.<sup>11</sup> Thus, we imagine the central bank in each period  $t - 1$  announcing a

---

<sup>9</sup> The interpretation of loss functions in models in the Barro-Gordon framework is always subtle and complicated. One can construct (mostly from arguments given elsewhere in this literature) a justification of (2.4) as a true social loss function. In short, the variable  $z_t$  is the desired rate of employment, which differs from the natural rate due to time-varying distortions about which the central bank has special knowledge. More in line with our preferred interpretation, one can arrive at both (2.4) and (2.7) as different aggregations of heterogeneous individual losses with (2.7) involving more representative weights. We prefer to interpret them less literally as approximations that capture essential features of the problem.

<sup>10</sup> In all the equilibria below, this variable is a sufficient statistic for what the history of policy outcomes has taught the public about the central bank's preferences.

<sup>11</sup> As noted below, the welfare-maximizing policy involves a one-period-ahead inflation expectation of zero;

zero inflation target for period  $t$ , and we measure the credibility,  $c_{t-1}$ , of that announcement as minus the absolute value of the public's inflation expectation:

$$c_{t-1} \equiv -|\pi_t|_{t-1}. \quad (2.8)$$

The further inflation expectations are from zero, the lower is credibility.<sup>12</sup>

Whether or not the reputation correctly characterizes the bank's preferences is almost as important as reputation itself. When Prell referred to a "credibility bind" in the FOMC meeting, he probably did not simply believe that the Fed's reputation was for allowing inflation to go above 5 percent, he also thought that this view did not accurately reflect the Fed's low-inflation zeal. Thus, not only bad reputation, but also misperceived bad reputation is important. In our model, this situation will correspond to the case when  $z_{t|t-1}$  is greater than  $z_t$  (and  $z_t$  is positive).

*Transparency* has to do with how easy it is for the public to deduce central bank intentions from observables. When  $z_t$  is not observable, the public learns about  $z_t$  from the economic outcomes, but the observable outcomes ( $\pi_t$ ,  $\varepsilon_t$  and  $l_t$ ) do not perfectly reveal the central bank's intentions due to the control error,  $\eta_t$ . Reducing the control error variance,  $\sigma_\eta^2$ , makes deducing  $z_t$  easier, and thereby, increases transparency. This is the only notion of varying transparency in CM, and this notion has the unfortunate limitation that increasing transparency simultaneously has the direct benefit of improving the control of the central bank.

To clarify the distinction between the control error and transparency, we assume that the control error satisfies

$$\eta_t \equiv \xi_t + \nu_t, \quad (2.9)$$

where  $\xi_t$  and  $\nu_t$  are independent mean-zero normal shocks. The private sector observes  $\xi_t$  at the end of period  $t$ , whereas the component  $\nu_t$  remains unobservable. The variances of  $\xi_t$  and  $\nu_t$  fulfill

$$\begin{aligned} \sigma_\xi^2 &= \tau \sigma_\eta^2 \\ \sigma_\nu^2 &= (1 - \tau) \sigma_\eta^2, \end{aligned} \quad (2.10)$$

---

then inflation at  $t$  responds only to the supply shock at  $t$ .

<sup>12</sup> CM ([7], p. 1108) define credibility in a different way: "Credibility is defined as the absolute value of the difference between the policymaker's plans and the private sector's beliefs about those plans. The smaller this difference, the higher the credibility of planned monetary policy." This corresponds to  $-|i_t - i_{t|t-1}|$  in our notation. In this view, the public does not know whether or not the bank is credible, whereas in our view, the credibility of an announcement is something judged by those to whom the announcement is made. Further, in our view, it is possible that a bank might credibly announce a zero-inflation policy when intending to implement higher inflation. Thus, we believe that our definition of credibility is closer to the common usage of the term.

where we call the parameter  $\tau$  ( $0 \leq \tau \leq 1$ ) (the degree of) transparency. Thus,  $\tau$  measures the share of the control-error variance that arises from the observable component. For  $\tau = 0$ , “minimum transparency,” no part of the control error is observed by the public. For  $\tau = 1$ , “full transparency of intention,” the whole control error is observable and the public can perfectly infer the bank’s intention from  $\pi_t$ .

The paradigm case of increased transparency is probably the immediate release of FOMC transcripts. This would not *directly* alter monetary control, but does *ceteris paribus* make it easier for the public to deduce the Fed’s intentions. Similarly, in inflation-targeting countries, the regular publication of informative Inflation Reports or Monetary Policy Statements makes it easier for the public to deduce the central bank’s intentions.

### 2.3 Three regimes

In order to study the roles of credibility and transparency, we focus on three monetary policy regimes, which differ by the degree of transparency, but have in common a lack of a commitment technology, so the central bank minimizes its loss function (2.3) under discretion. These are:

- U** Unobservable goal and intention: In this regime,  $0 \leq \tau < 1$ , and  $z_t$  and  $i_t$  are not observed by the private sector. In period  $t$ , the private sector observes only  $\pi_t$ ,  $l_t$ ,  $\xi_t$  and  $\varepsilon_t$ .
- OI** Observable intention: This is regime U with  $\tau = 1$ , full transparency of *intention*. The private sector does not observe  $z_t$  directly, but it observes  $\pi_t$ ,  $l_t$ ,  $\varepsilon_t$ , and  $\eta_t$ , from which it can deduce  $i_t$  and, in equilibrium,  $z_t$ , without error.
- OG** Observable goal: “Extreme” transparency. In period  $t$ , the private sector directly observes  $z_t$ ,  $\pi_t$ ,  $\eta_t$ ,  $\varepsilon_t$ , and  $l_t$ .

Regime U is our baseline case. Regime OI is the limit of regime U when transparency reaches its maximum. We will show that the public can infer the bank’s goal perfectly in regime OI, but the equilibrium is remarkably different from the equilibrium in regime OG in which the goal is directly observed rather than perfectly inferred.

As a basis of comparison, we consider regime S (the social optimum) in which the central bank is forced to commit to a policy rule that minimizes the social loss function, (2.3) with (2.7). This results in the standard commitment solution,

$$i_t = -\frac{1}{2}\varepsilon_t. \tag{2.11}$$

The policy then optimally smooths the effect of the supply shock between inflation and employment. The policy obviously disregards  $z_t$ , which does not enter the social loss function. The regime studied by Cukierman-Meltzer [7] is reviewed in Appendix D.

It is relevant to ask why the other regimes are of interest when the simple rule (2.11) could be imposed to maximize social welfare. We believe that, in the real world, policy under discretion arises because the complexity of the economic and political environment make codification, adoption, and verification of a good policy rule impossible. In any formal model that can be solved, a forcing rule may seem like the obvious answer. We nevertheless believe that studying discretion and transparency in solvable models is a useful way to gain insights about more complex environments in which analytical results cannot be obtained.<sup>13</sup>

## 2.4 Generic economic dynamics for all regimes

The analysis of these regimes is greatly simplified by the fact that the dynamics of the economy, up to the parameters of the central bank policy rule, are the same in each regime. In all regimes, we assume that the private sector believes that the central bank's policy follows

$$i_t = k_0 + k_1\varepsilon_t + k_2z_t + k_3z_{t|t-1}, \quad (2.12)$$

for some coefficients  $k_0, \dots, k_3$ .<sup>14</sup>

We confirm in section 3 that, if the private sector believes the policy is given by (2.12), the central bank will optimally behave according to (2.12). This assumption has the effect of making a simple linear learning scheme optimal for the private sector and, in particular, rules out signalling equilibria in which small changes in policy can signal sharp differences in central bank type.

Given the private sector's belief in (2.12), expected inflation is given by

$$\pi_{t|t-1} = k_0 + (k_2 + k_3)z_{t|t-1}, \quad (2.13)$$

and employment evolves according to

$$l_t = i_t + \eta_t - k_0 - (k_2 + k_3)z_{t|t-1} + \varepsilon_t. \quad (2.14)$$

These expressions will hold whether or not the private sector's beliefs about policy are rational. In rational expectations equilibrium, the central bank behaves according to (2.12), and

<sup>13</sup> McCallum [22] has criticized the commitment and discretion framework with the argument that the discretion equilibrium is irrelevant, since central banks would simply see through the problem and do the right thing. We take up this argument in the conclusions.

<sup>14</sup> The function (2.12) is a simplification of

$$i_t = k_0 + k_1\varepsilon_t + k_2z_t + k_3z_{t|t-1} + k_4z_{t-1},$$

which allows policy to depend separately on  $z_t$  and  $z_{t-1}$ , and hence separately on  $\theta_t$ . None of our results change if we begin by assuming that the public believes policy is made according to (??). That is, for all the cases we consider, the equilibrium decision rule ends up having  $k_4 = 0$ .

equilibrium dynamics are,

$$\pi_t = k_0 + k_1\varepsilon_t + k_2z_t + k_3z_{t|t-1} + \eta_t \quad (2.15)$$

$$\pi_{t|t-1} = k_0 + (k_2 + k_3)z_{t|t-1} \quad (2.16)$$

$$\pi_t - \pi_{t|t-1} = k_1\varepsilon_t + k_2(z_t - z_{t|t-1}) + \eta_t \quad (2.17)$$

$$l_t = (1 + k_1)\varepsilon_t + k_2(z_t - z_{t|t-1}) + \eta_t \quad (2.18)$$

$$l_t - l_t^* = (1 + k_1)\varepsilon_t + k_2(z_t - z_{t|t-1}) + \eta_t - l^* - z_t \quad (2.19)$$

The only endogenous variable not determined here is the key to the analysis: reputation,  $z_{t|t-1}$ . The next section completes the derivation of the rational-expectations equilibria for the various regimes.

### 3 Solving the model

#### 3.1 Regime U: unobservable goals and intentions

CM solve the model for their case by solving the private sector's learning and optimization problem and the central bank's optimization problem directly. This solution is complicated and becomes intractable when considering the more standard loss function we use. We find a solution by noting that the Kalman filter provides the optimal solution to the public learning problem and by casting the central bank optimization as a dynamic programming problem. Of course, for the CM regime, our approach gives the same solution (Appendix D).

We first derive the public's learning rule about  $z_t$ , and then the optimal  $k$ s in the policy function. Since the public does not observe  $z_t$  or  $i_t$  directly, it forms its expectation of inflation for period  $t$  at the end of period  $t - 1$  based only on the history of  $\pi_t$ ,  $l_t$  and  $\xi_t$ . At the end of period  $t$ , the public can construct the variable

$$y_t \equiv \pi_t - k_0 - k_1\varepsilon_t - k_3z_{t|t-1} - \xi_t = i_t + \nu_t - k_0 - k_1\varepsilon_t - k_3z_{t|t-1}, \quad (3.1)$$

where we have used (2.9). Under the public's assumption that policy is made according to (2.12), we have

$$y_t = k_2z_t + \nu_t. \quad (3.2)$$

Further, under (2.12),  $y_t$  contains all the new private-sector information about  $z_t$  that arrives in period  $t$ :  $E^P[z_t|y_t, z_{t|t-1}] = E^P[z_t|\text{all private-sector information in period } t]$ .<sup>15</sup> Believing that

<sup>15</sup> This is easily verified, assuming that  $z_{t|t-1} = E^P[z_t|\text{all private sector information in period } t - 1]$ .

it observes  $k_2 z_t$  plus a normal error, the private sector's learning problem is optimally solved using the Kalman filter, treating (2.6) as the transition equation and (3.2) as the measurement equation. The steady-state solution to this problem gives the dynamics of reputation:<sup>16</sup>

$$z_{t+1|t} = (\rho - gk_2) z_{t|t-1} + gy_t, \quad (3.3)$$

where  $g$  is the Kalman gain and can be expressed in terms of  $k_2$  and the exogenous parameters only.<sup>17</sup>

Under the private sector's belief (2.12),  $\pi_t$ ,  $l_t$ ,  $z_t$  and  $z_{t|t-1}$  evolve as in (2.2), (2.14), (2.6), and (3.3), respectively. There are two state variables in this economy, and for our purposes, it is natural to take the employment target,  $z_t$ , and reputation,  $z_{t+1|t}$  as state variables.

To solve for the optimal policy rule for  $i_t$ , we recursively define the central bank's (steady-state) value function as

$$V(z_{t|t-1}, z_{t-1}) \equiv \mathbb{E}_{t-1} \min_{i_t} \mathbb{E}_{t-} \left[ L_t + \beta V(z_{t+1|t}, z_t) \right], \quad (3.4)$$

where  $\mathbb{E}_{t-}$  denotes the expectation of the central bank given its information at the *beginning* of period  $t$ , after it has observed  $\varepsilon_t$  and  $\theta_t$ , but before  $\eta_t$ ,  $\pi_t$ , and  $l_t$  have been realized. Because the loss function is quadratic and the two state variables evolve linearly (according to (2.6) and (3.3)), the value function is quadratic,

$$V(z_{t|t-1}, z_{t-1}) = \delta_0 + \delta_1 z_{t|t-1} + \frac{1}{2} \delta_2 z_{t|t-1}^2 + \delta_3 z_{t-1} + \frac{1}{2} \delta_4 z_{t-1}^2 + \delta_5 z_{t|t-1} z_{t-1}, \quad (3.5)$$

where the coefficients  $\delta_0, \dots, \delta_5$  remain to be determined.

In period  $t$ , the central bank's problem is then to solve

$$\min_{i_t} \mathbb{E}_{t-} \left[ L_t + \beta V(z_{t+1|t}, z_t) \right]. \quad (3.6)$$

The first-order condition with respect to  $i_t$  is

$$i_t + \mathbb{E}_{t-} l_t - l^* - z_t + \beta \mathbb{E}_{t-} \left[ (\delta_1 + \delta_2 z_{t+1|t} + \delta_5 z_t) \frac{\partial z_{t+1|t}}{\partial i_t} \right] = 0, \quad (3.7)$$

where the derivative  $\frac{\partial z_{t+1|t}}{\partial i_t}$  enters because current policy affects future reputation through (3.1) and (3.3). The expectations and the partial derivative in this expression can be evaluated using

<sup>16</sup> That is, when the forecast error variance has converged. See appendix A.

<sup>17</sup> CM's first definition of credibility was discussed above. CM second definition is given on p. 1109: " $\lambda$  [which equals our  $\rho - gk_2$ ] can therefore be taken as a measure of credibility. The higher  $\lambda$ , the longer it takes the public to recognize a change in governmental objectives and the lower, therefore, the government's credibility." We prefer to interpret  $\rho - gk_2$  as (one factor in) the persistence of reputation.

expressions already shown, and the resulting expression can be solved for  $i_t$ , obtaining a policy rule of the form (2.12) with coefficients (see appendix B)

$$k_0 = l^* - \beta g \delta_1 \quad (3.8)$$

$$k_1 = -\frac{1}{2} \quad (3.9)$$

$$k_2 = \frac{1 - \beta g \delta_5}{2 + \beta g^2 \delta_2} \quad (3.10)$$

$$k_3 = k_2 - \beta g (\rho - g k_2) \delta_2. \quad (3.11)$$

In appendix B, we show that solving for the  $k$ s can be reduced to solving a single nonlinear equation,  $k_2 = f(k_2)$ , where  $f(k_2) \in [0, \frac{1}{2}]$  for all  $k_2$  and  $f$  is continuous on  $[0, \frac{1}{2}]$ . A simple fixed-point argument guarantees the existence of at least one solution. Thus, there is an equilibrium under discretion with a policy rule (a “time-consistent” policy rule) of the form assumed by the private sector in (2.12).<sup>18</sup> While we have not proved uniqueness of the solution, we have confirmed uniqueness numerically, as discussed below.

To highlight some crucial features of transparency and credibility we now examine regime OI (observable instrument), which is the limit of case U as transparency of intentions reaches its upper limit, and case OG (observable goal) in which the central bank’s goal is directly observable.

### 3.2 Regime OI: Observable Intentions

The solution in regime OI is obtained from that of regime U by assuming full transparency of intentions ( $\tau = 1$ ). Thus, the  $k$ s follow by letting  $\sigma_v^2$  go to zero in the expressions for the baseline regime, (3.8)–(3.11) and (B.9)–(B.13). Taking the relevant limits gives,

$$k_0 = \frac{1 - \beta \rho}{1 + \beta \rho} l^* < l^* \quad (3.12)$$

$$k_1 = -\frac{1}{2} \quad (3.13)$$

$$k_2 = k_3 = \frac{1 - \beta \rho^2}{2(1 + \beta \rho^2)} < \frac{1}{2}, \quad (3.14)$$

where we have used that  $gk_2 = \rho$  when  $\sigma_v^2 = 0$  (see Appendix C).

---

<sup>18</sup> It is worth emphasizing that the derivation rests on the assumption that the private sector assumes the central bank acts according to (2.12). There are almost surely some other equilibria of the model without this assumption. CM implicitly make the same assumption, and Rogoff [26] pointed out the likely existence of equilibria without the assumption.



### 3.3 Regime OG: Observable Goal

In regime OG, we allow the private sector to observe  $z_t$  directly in period  $t$ . Thus,

$$z_{t|t-1} \equiv \rho z_{t-1} \quad (3.15)$$

independently of the policy rule. The value function (3.4) from the baseline regime is still appropriate, but with (3.15). Inflation expectations,  $\pi_{t|t-1}$ , are given by (2.16) with (3.15),

$$\pi_{t|t-1} = k_0 + (k_2 + k_3)\rho z_{t-1}. \quad (3.16)$$

Since  $\frac{\partial z_{t+1|t}}{\partial i_t} \equiv 0$ , the first-order condition with respect to  $i_t$ , (3.7), is now

$$E_t \left[ (i_t + \eta_t) + (i_t + \eta_t - \pi_{t|t-1} + \varepsilon_t - l^* - z_t) \right] = 2i_t - \pi_{t|t-1} + \varepsilon_t - l^* - z_t = 0,$$

which with (3.16) implies

$$i_t = \frac{1}{2}(k_0 + l^*) + \frac{1}{2}\varepsilon_t + \frac{1}{2}z_t + \frac{1}{2}(k_2 + k_3)\rho z_{t-1} + \frac{1}{2}\theta_t.$$

This is of the form (2.12) with the coefficients

$$k_0 = l^* \quad (3.17)$$

$$k_1 = -\frac{1}{2} \quad (3.18)$$

$$k_2 = k_3 = \frac{1}{2}. \quad (3.19)$$

### 3.4 Numerical analysis of the model

We summarize the numerical approach here; for details, see Appendices B and E. In order to explore properties that could not be proved analytically for all parameter values, we study the properties for a large number of particular parameter values. Specifically, we solve the model for 100,000 points drawn uniformly from the parameter space  $(\beta, \rho, \tau) \in [0, 1]^3$ ,  $(\sigma_\eta^2, \sigma_\theta^2, \sigma_\varepsilon^2, l^*) \in [0, 10]^4$ . Once the model is solved, we tally which among a large number of claims hold true for that parameter value, e.g.: Is central bank loss in regime U lower than in OI? Is the derivative of the central bank loss with respect to transparency positive?

If a property holds for some parameter values and not for others, the solutions for the particular parameter values constitute a constructive proof that the result is indeterminate. If a property holds for all 100,000 points, we do not have proof that the property holds for all values, but one can make a very strong probabilistic statement. If a claim holds for each of  $N$  draws,

Table 3.1. Summary of regimes

Regime	$k_0$	$k_1$	$k_2$	$k_3$	$\rho - gk_2$	$P$	$s$
U	$< l^*$	$-\frac{1}{2}$	$< \frac{1}{2}$	$< k_2$	$< \rho$	$> \sigma_\theta^2$	$> 1$
OI	$< l^*$	$-\frac{1}{2}$	$< \frac{1}{2}$	$= k_2$	0	$\sigma_\theta^2$	$\infty$
OG	$l^*$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$			
S	0	$-\frac{1}{2}$	0	0			

Notes:  $P$  denotes the variance of the forecast error,  $E[(z_t - z_{t|t-1})^2]$ .

the probability that the claim is false on a fraction of the parameter space of at least size  $\omega$  is less than or equal to  $(1 - \omega)^N$ . Thus, with 100,000 draws, the probability that the claim is false for at least 0.01 percent of the parameter space is less than 0.005 percent.<sup>19 20</sup>

### 3.5 Summary of regimes U, OI and OG

In order to facilitate discussion, we restate the equations for the dynamics of the economy that are common across regimes U, OI, and OG:

$$\pi_t = k_0 + k_2 z_t + k_3 z_{t|t-1} - \frac{1}{2} \varepsilon_t + \eta_t \quad (3.20)$$

$$\pi_{t|t-1} = k_0 + (k_2 + k_3) z_{t|t-1} \quad (3.21)$$

$$\pi_t - \pi_{t|t-1} = k_2(z_t - z_{t|t-1}) - \frac{1}{2} \varepsilon_t + \eta_t \quad (3.22)$$

$$l_t = k_2(z_t - z_{t|t-1}) + \frac{1}{2} \varepsilon_t + \eta_t \quad (3.23)$$

$$z_{t+1|t} = (\rho - gk_2) z_{t|t-1} + g(k_2 z_t + \nu_t) \quad (3.24)$$

$$= \rho \left[ \frac{1}{s} z_{t|t-1} + \left(1 - \frac{1}{s}\right) z_t \right] + g\nu_t, \quad (3.25)$$

where in (3.25) we observe that the dynamics of reputation can also be expressed in terms of a signal-to-noise ratio,  $s$ , defined by the ratio of the variances of the signal in the Kalman filter,  $y_t = k_2(z_t - z_{t|t-1}) + \nu_t$ , and the noise,  $\nu_t$ ,

$$s \equiv \frac{k_2^2 \text{Var}[z_t - z_{t|t-1}] + \sigma_\nu^2}{\sigma_\nu^2} \geq 1 \quad (3.26)$$

(see appendix A for details). Some properties of key coefficients can be shown analytically and are summarized in Table 3.1.

The next two sections take up the comparative behavior of banks with different levels of credibility and of transparency, respectively.

<sup>19</sup>  $0.9999^{100,000} \approx 0.000045$ .

<sup>20</sup> Rather than using a uniform draw from the parameter space, with a meaningful *prior* density measuring the empirical relevance of various regions, one can produce more meaningful *posterior* measures of the empirical relevance of the computed properties.

## 4 Credibility and optimal policy

Our regime U is, perhaps, the most straightforward generalization of the Barro-Gordon model [3] that gives rich dynamics in reputation and credibility. The Barro-Gordon model is a special case of our OG regime in which the employment target is constant ( $\rho = \sigma_\theta^2 = 0$ ) and there is no control error ( $\sigma_\nu^2 = \sigma_\eta^2 = 0$ ).<sup>21</sup> Imposing these parameter constraints in our OG solution gives the Barro-Gordon solution,  $\pi_t = i_t = l^* - \frac{1}{2}\varepsilon_t$  and  $l_t = \frac{1}{2}\varepsilon_t$ . In this equilibrium, if the central bank announced the welfare-maximizing, zero-inflation policy, it would not be credible,  $c_t \equiv -l^*$ . Credibility is not stochastic, however, and this model offers few insights about the sort of dynamic issues captured in the folk wisdom discussed in the introduction.

In regime U, stochastic and private central-bank goals combined with noise that masks the bank's true intentions lead to an equilibrium with time-varying credibility. Then credibility (using (2.8) and (3.21)) is given by

$$c_t = -|k_0 + (k_2 + k_3)z_{t|t-1}|,$$

with dynamics implied by those of reputation,  $z_{t|t-1}$ , given by (3.25). The persistence of reputation depends positively on the persistence of the central bank's employment target,  $\rho$ , and negatively on the signal-to-noise ratio,  $s$ . Starting with a correctly perceived central bank ( $z_{t|t-1} = \rho z_{t-1}$ ), a positive employment target shock in period  $t$  ( $\theta_t > 0$ ) pushes  $z_t$  above the reputation. By (3.22) the bank, everything else equal, will choose a policy that is *more* inflationary than expected by the public. A positive unobservable component of the control error ( $\nu_t > 0$ ) will make inflation at  $t$  higher than the bank intended, by (3.25) erroneously pushing up the bank's inflationary reputation. Policy at  $t + 1$  will then, everything else equal, be *less* inflationary than expected. The remainder of this section discusses claims (1) and (2) in the introduction: whether, for given transparency, a low-credibility bank will pursue a more or less expansionary monetary policy than a high-credibility bank, and whether a low-credibility bank responds differently to a supply shock.

### 4.1 Credibility and optimal policy with fixed transparency

The first claim in the introduction is that a central bank with low credibility may optimally need to follow a less expansionary inflation policy—sacrificing some employment—in order to regain

---

<sup>21</sup> Barro and Gordon [3] also include the case of an observable exogenous autoregressive natural rate of unemployment, which is similar to our regime OG.

credibility, than would a bank with high credibility. The second claim is that a low-credibility bank has less flexibility optimally to respond to supply shocks. The first claim only holds under careful interpretation in our model; the second does not hold.

To assess these claims, we compare two realizations of the economy, starting in the beginning of period  $t$  with the same value of the state variable  $z_{t-1}$ , but in one case credibility is low ( $\ell$ ) and in the other it is high ( $h$ ),  $c_{t-1}^\ell < c_{t-1}^h$ . Credibility alone does not tell the sign of inflation expectations; we restrict the discussion of a low and a high credibility bank to situations of positive (that is, too high) inflation expectations:  $\pi_{t|t-1}^\ell > \pi_{t|t-1}^h \geq 0$ . By (3.21), the two banks' reputations will then fulfill  $z_{t|t-1}^\ell > z_{t|t-1}^h$ . For any variable  $x_t$ , define  $\Delta x_t \equiv x_t^\ell - x_t^h$ , the low-credibility value minus the high-credibility value. We thus have

$$\Delta\pi_{t|t-1} > 0, \Delta z_{t|t-1} > 0. \quad (4.1)$$

The following proposition summarizes our results:

**Proposition 4.1.** *In regime U, ceteris paribus:*

- (i) *The low-credibility bank optimally implements higher inflation than the high-credibility bank,  $\Delta\pi_t > 0$ ;*
- (ii) *The low-credibility bank optimally implements lower inflation relative to private sector expectations  $\Delta(\pi_t - \pi_{t|t-1}) < 0$ . This larger negative inflation surprise leads to lower employment in the low-credibility economy,  $\Delta l_t < 0$ ; and*
- (iii) *The low-credibility bank responds to supply shocks and to shocks to the employment goal in the same way as does the high-credibility bank.*

Part (i) and (ii) follow directly from (3.20), (3.21), (3.23) and (4.1), since

$$\Delta\pi_t = k_3\Delta z_{t|t-1} < (k_2 + k_3)\Delta z_{t|t-1} = \Delta\pi_{t|t-1}$$

and  $k_3 < k_2 + k_3$ . The low-credibility bank accommodates part of, but only part of, the higher inflation expectations it faces, resulting in higher inflation. The negative inflation surprise is larger in absolute terms under low credibility, leading to lower employment,

$$\Delta l_t = \Delta(\pi_t - \pi_{t|t-1}) = -k_2\Delta z_{t|t-1} < 0.$$

The low-credibility bank would, of course, gain reputation and credibility faster if it accommodated less of the inflation expectation, but this is not optimal, due to the current employment cost of doing so. Re-stating the first-order condition, (3.7), gives

$$\mathbf{E}_t - \frac{1}{2} \frac{\partial \pi_t^2}{\partial i_t} + \mathbf{E}_t - \frac{1}{2} \frac{\partial (l_t - l^* - z_t)^2}{\partial i_t} + \beta \mathbf{E}_t - \frac{\partial V(z_{t+1|t}, z_t)}{\partial z_{t+1|t}} \frac{\partial z_{t+1|t}}{\partial i_t} = 0. \quad (4.2)$$

Lowering  $i_t$  leads to a larger negative inflation surprise, yielding benefits in terms of lower inflation and better future reputation (the first and third terms). The negative surprise lowers employment, which increases the loss (the second term). The optimal policy is a compromise between these concerns. This result is one formalization of results from the “gradualist versus cold turkey” debate regarding lowering inflation in the early 1980s.<sup>22</sup> While the optimal speed of adjustment will vary depending on the model, the result that is likely to generalize is that a bank facing low credibility should move more slowly, not more quickly than one facing high credibility.

Given this result, evidence of higher inflation and higher inflation expectations is not necessarily evidence that a bank is insufficiently attentive of its inflation target; rather, it may be evidence of low credibility to which the bank is optimally responding. Only evidence that the low-credibility bank is implementing smaller absolute inflation surprises is evidence that it is behaving suboptimally.

With regard to part (iii), it is clear that the low-credibility and high-credibility banks respond in precisely the same way to the supply shock,  $\varepsilon_t$ , and to the employment target shock,  $\theta_t$ . Thus, low-credibility banks do not have less flexibility with which to respond to supply shocks than high-credibility banks. Further, banks do not build up credibility so as to spend it (disproportionately) when the employment target is highest. Both results stem from the linear nature of our model. We believe, however, that these results form an important baseline: conflicting results in a model like this must rest on important nonlinearities.

#### 4.2 An exception: A patient bank with very persistent goals

The specific results above and below about the effects of credibility and transparency do not follow in the special case of no discounting and a very persistent employment target. In the limit when  $\beta\rho \rightarrow 1$ , the socially optimal policy rule (regime S) is the solution for both regimes U and OI.

**Proposition 4.2.** *In regimes U and OI, in the limit as the persistence of the idiosyncratic goals of the bank ( $\rho$ ) and the central bank’s discount factor ( $\beta$ ) both go to one, the socially optimal outcome results.*

For regime OI, it follows from (3.12)–(3.14) that in the limit,  $k_0 = k_2 = k_3 = 0$ , which is the policy rule in regime S, (2.11). The result is only slightly more difficult to see for regime U (Appendix B).

---

<sup>22</sup> See, e.g., Fuhrer [13] and Ball [2].

The intuition for this result is as follows. With  $\rho$  near one, the private sector believes that any shock to the employment target will be long-lasting. Thus, if an inflationary action increases  $z_{t+1|t}$ , this increase is very persistent. With  $\beta$  near one, future costs of this increase weigh heavily on current decisions by the bank. By setting  $k_2 = 0$ , the bank guarantees that the private sector will not attribute any inflation surprise to an increase in the employment target. Thus, changes in the reputation become so costly for the bank with these parameters, that it prefers to give up any attempt to follow its idiosyncratic goals. The bank's concern about its reputation becomes so strong that it simply chooses to pursue the socially preferred policy.

## 5 The effects of transparency

Next we examine statement (3) in the introduction: does increased transparency improve credibility and the policy outcome? First, we consider the effects of increased transparency. Second, we consider the difference between observing the goal directly (regime OG, full transparency of goals) and inferring it in equilibrium (regime OI, full transparency of intention). We show that extreme transparency, as represented by directly observable central bank goals, is the worst of all regimes in our model. Third, we examine a numerical example for typical parameters. Fourth, we examine the optimal degree of transparency in regime U. Finally, we show that with a patient central bank without average inflation bias, increased transparency is always better for society but worse for the bank.

To judge policy outcomes, we report results on “welfare” as measured by the unconditional expectation of the relevant loss function. Thus, we learn which regime is best on average, or which would be preferred without knowledge of the state variables. The social unconditional loss is proportional to  $E[L_t^p]$  with  $L_t^p$  given by (2.7), which can be written as

$$E[L_t^p] = \frac{1}{2} \left( k_0^2 + \text{Var}[\pi_t] + \text{Var}[l_t] + l^{*2} \right). \quad (5.1)$$

The central bank's unconditional loss is proportional to  $E[L_t]$ , with  $L_t$  as in (2.4), which can be written as a sum of six terms:

$$E[L_t] = \frac{1}{2} \left( k_0^2 + \text{Var}[\pi_t] + \text{Var}[l_t] + l^{*2} + \text{Var}[z_t] - 2\text{Cov}[l_t, z_t] \right). \quad (5.2)$$

The central bank loss differs from the social loss by the term  $\frac{1}{2}(\text{Var}[z_t] - 2\text{Cov}[l_t, z_t])$ , where only the covariance term is endogenous. The intuition for this is that the central bank optimum differs from the private sector optimum only to the extent that the central bank can generate

movements in employment that follow the target  $z_t$ . The central bank can do this only by using inflation surprises.

## 5.1 Increasing transparency in case U

Our results are summarized as follows:

**Proposition 5.1.** *Consider raising  $\tau$  in regime U.*

(i) Reputation. *Raising  $\tau$  raises the variance of reputation, but decreases the variance of reputational errors,  $(z_t - z_{t|t-1})$ . Raising  $\tau$  raises the sensitivity of reputation to the central bank's intention (increases  $\partial z_{t+1|t}/\partial i_t$ ).*

(ii) Inflation. *Raising  $\tau$  lowers average inflation (strictly whenever  $l^* > 0$ ), but may raise or lower the variance of inflation and the inflation term in the unconditional loss,  $E[\pi_t^2]$ .*

(iii) Employment. *Raising  $\tau$  lowers the variance of employment and lowers the employment term in the social unconditional loss,  $E[(l_t - l^*)^2]$ , but raises the employment term of the central bank unconditional loss,  $E[(l_t - l_t^*)^2]$ .*

(iv) Unconditional loss. *Raising  $\tau$  may raise or lower social and central bank unconditional loss. For plausible discount factors ( $\beta > 0.5$ ), social loss always falls.*

Each part of the claim is about the derivative of some equilibrium value with respect to  $\tau$ . Part (i) is shown analytically. Parts (ii) through (iv) are demonstrated numerically as discussed in Section 3.4.

Part (i), reputation. If raising  $\tau$  deserves the interpretation as increasing transparency, then the variance of reputational errors should fall. Perhaps not as obvious, the variance of reputation should rise, since reputation,  $z_{t|t-1}$  will more closely track the actual goal,  $z_t$ . This is indeed the case. The noise in the Kalman filter ( $\sigma_\nu^2 = (1 - \tau)\sigma_\eta^2$ ) falls, raising the signal-to-noise ratio,  $s$ , and the Kalman gain,  $g$  (see appendix A). The private sector obtains a better estimate of the employment target:  $\text{Var}[z_t - z_{t|t-1}] \equiv P$  falls. By the same reasoning, the unconditional variance of reputation ( $\text{Var}[z_{t|t-1}] = \text{Var}[z_t] - P$ ) rises, as the predictor better tracks the predicted. The rise in the Kalman gain increases the sensitivity of  $z_{t+1|t}$  to  $i_t$ , since  $\partial z_{t+1|t}/\partial i_t = g$ .

The remaining parts of the proposition are shown numerically, but some intuition can be derived from noting how increases in transparency affect the optimal  $k$ s in the policy rule. Of course,  $k_1 = -\frac{1}{2}$  in all cases. Numerically, the remaining  $k$ s fall with rises in  $\tau$  for all parameter values.<sup>23</sup>

Consider why  $k_0$ , average inflation, falls. In any Barro-Gordon-type model, inflation must be high enough on average to keep the central bank from engineering positive inflation surprises on average. Thus, the fall in  $k_0$  must reflect the fact that greater transparency reduces the

<sup>23</sup>The  $k$ s fall strictly, except at extreme values for  $\beta$ ,  $\rho$ , and  $l^*$ .

Table 5.1. An increase in the degree of transparency,  $\tau$

Parm. space	$s$	$P$	$g$	Var[ $x_t$ ]		E[ $x_t^2$ ]			Loss	
				$z_{t t-1}$	$\pi_t$	$\pi_t$	$l_t - l_t^*$	$l_t$	E[ $L_t$ ]	E[ $L_t^p$ ]
Full	+	-	+	+	43.8	93.1	+	-	79.6	96.3
Small	+	-	+	+	87.5	87.5	+	-	+	-

Notes: Plus and minus indicate unambiguous signs of the derivate with respect to  $\tau$ . Numbers indicate the proportion of the parameter space for which the sign is negative. The “full” parameter space refers to the parameter space  $(\beta, \rho, \tau) \in [0, 1]^3$ ,  $(\sigma_\eta^2, \sigma_\varepsilon^2, \sigma_\theta^2, l^*) \in [0, 10]^4$ . The “small” parameter space is the same except that  $\beta = 0.99999$  and  $l^* = 0$ .

(average) net benefit of inflation surprises to the bank. To see that this is the case, start from a fixed intended rate of inflation,  $i_t$ . If the bank chooses a more inflationary policy, this leads to higher employment today (which, with  $l^* > 0$ , is good on average) but also worsens the bank’s inflationary reputation by raising  $z_{t+1|t}$ :  $\partial z_{t+1|t} / \partial i_t = g > 0$ . With greater transparency, the marginal degradation of reputation is larger, since  $g$  is larger. This greater reputational effect has costs in terms of higher future optimal inflation and lower future employment—the third term in (4.2). Thus, transparency raises the reputational cost of cheating, allowing a lower average inflation.

The argument for why  $k_2 + k_3$  falls when transparency rises is very similar to the argument for  $k_0$ ; the argument for the unconditional expectation of inflation is merely restated for the inflation expectation conditional on a given level of  $z_{t|t-1}$ . When reputation is  $z_{t|t-1}$ , the public expects an inflation rate of  $k_0 + (k_2 + k_3)z_{t|t-1}$ . This expected inflation must be high enough to prevent the central bank from implementing surprises on average when its reputation is  $z_{t|t-1}$ . The marginal cost to the central bank of inflation surprises due to the reputation effect rises linearly with the level of  $z_{t|t-1}$ . When the effect is magnified by raising  $\tau$  and, hence,  $g$ , a lower  $(k_2 + k_3)$  will be sufficient to deter the bank from surprises. The arguments for  $k_2$  and  $k_3$  separately are similar.

Overall, increased transparency makes the bank less activist, by reducing  $k_0$ ,  $k_2$ , and  $k_3$ . Results (ii) through (iv) of the proposition follow relatively directly from this fall in activism. Some further insight is possible from looking at changes in key quantities (Table 5.1, the row for the “full” parameter space,  $(\beta, \rho, \tau) \in [0, 1]^3$ ,  $(\sigma_\eta^2, \sigma_\varepsilon^2, \sigma_\theta^2, l^*) \in [0, 10]^4$ ).

Part (ii), inflation. The fall in average inflation was discussed above. From (3.20), the unconditional variance of inflation is

$$\text{Var}[\pi_t] = k_2^2 \text{Var}[z_t] + [k_3^2 + 2(k_2 + k_3)] \text{Var}[z_{t|t-1}] + \frac{1}{4} \sigma_\varepsilon^2 + \sigma_\eta^2, \quad (5.3)$$



where we have used that  $\text{Cov}[z_t, z_{t|t-1}] = \text{Var}[z_{t|t-1}]$ . The first term falls since  $k_2$  falls; the third and fourth are fixed. The change in the second term is ambiguous, since the coefficient within the bracket falls whereas the variance of reputation rises. For a significant portion ( $100 - 43.9 = 56.1$  percent) of the parameter space, the rise in the variance dominates, implying that the variance of inflation rises. The term  $k_0^2 + \text{Var}[\pi_t]$  in the loss function rises only when the increase in inflation variance dominates the fall in  $k_0^2$ , which is only on  $100 - 93 = 7$  percent of the parameter space.

Part (iii), employment. The employment term in the social loss function,  $\text{E}[(l_t - l^*)^2] = \text{Var}[l_t] + l^{*2}$ , always falls. The variance of employment falls because the variance of inflation surprises falls. In contrast, the employment term in the central bank loss function,  $\text{E}[(l_t - l_t^*)^2]$ , always rises. The intuition for this is that the central bank uses inflation surprises to generate a correlation between employment and the target,  $l_t^*$ . Greater transparency limits the ability to achieve this. Specifically, the fall in the variance of employment comes from the component of employment that is correlated with  $z_t$ .<sup>24</sup>

Part (iv), welfare. Given the results for the components of the loss functions, it is natural that both the central bank and social loss can either rise or fall with transparency. Social loss generally falls (96.3 percent of the parameter space) with increased transparency, however. Further, all of the cases in which social loss rises involve very low discount factors ( $\beta < 1/2$ ). For plausible discount rates, transparency is socially good.

Rises in transparency are also good for the central bank on 79.5 percent of the parameter space. It is important to note, however, that the loss rises for plausible parameter values, for example:  $\beta = 0.97, \rho = 0.30, \sigma_\eta = 1.89, \sigma_\theta = 1.0, \sigma_\varepsilon = 1.15, l^* = 0.11, \tau = 0.36$ . In this case, the target is moderately persistent and the control error has a standard deviation about twice that of the target shock and of the real shock.

---

<sup>24</sup> More formally, we have

$$\text{E}[(l_t - l_t^*)^2] = \text{Var}[l_t - z_t] + l^{*2} = \text{Var}[l_t] + l^{*2} + \text{Var}[z_t] - 2\text{Cov}[l_t, z_t].$$

The second and third terms on the right side do not change. The first term falls. It is, however, the component of  $l_t$  that covaries with  $z_t$  that is diminished in variance; thus, the fall in the covariance between  $l_t$  and  $z_t$  more than offsets the fall in variance of  $l_t$ :

$$\text{Var}[l_t] - 2\text{Cov}[l_t, z_t] = (k_2^2 - 2k_2)P + \frac{1}{4}\sigma_\varepsilon^2 + \sigma_\eta^2,$$

where we have used  $\text{Cov}[l_t, z_t] = k_2P$ . The left side rises, since  $P$  falls and since  $k_2^2 - 2k_2 < 0$  falls in magnitude.

## 5.2 Regime OG versus U and OI: observe the goal directly or infer it from actions?

In regimes U and OI, the public must attempt to deduce the bank's employment goal,  $z_t$ . This makes the bank's reputation sensitive to its intentions,  $\partial z_{t+1|t}/\partial i_t = g > 0$ : a positive inflation surprise today leads to a more inflationary reputation. This leads to costs to the bank given by the third term in (4.2). The reputation cost of inflation surprises remains even in regime OI when transparency of intention is at its maximum,  $\tau = 1$ . In this case, the public perfectly infers  $z_t$  in period  $t$  in equilibrium;  $z_{t|t-1} = \rho z_{t-1}$ ), but it remains true that if the central bank were to implement higher-than-equilibrium inflation, its reputation would suffer. This fact about off-the-equilibrium path behavior has important effects on the central bank's incentives.

In the OG regime, the public *directly* observes the bank's employment target and thus  $z_{t|t-1} \equiv \rho z_{t-1}$  independently of the bank's behavior, implying that  $\partial z_{t+1|t}/\partial i_t \equiv 0$ , and the third term in the first-order condition, (4.2), is zero. The absence of this reputational cost to its actions in case OG changes the incentives for the bank dramatically, leading to a striking result:

**Proposition 5.2.** *When the central bank's idiosyncratic goals are directly observed by the public (regime OG), average inflation, social loss, and central bank loss are each higher than under any level of transparency of intention with unobserved goal (regimes U and OI).*

This result is demonstrated numerically to hold as described in Section 3.4.

Thus, "extreme" transparency, in the sense that the public is no longer learning about the central bank's future intentions from current actions, is the worst of all worlds. Since the private sector has no illusions about the bank, the bank loses an important incentive to behave.

## 5.3 A numerical example

Table 5.2 presents a numerical example for typical parameter values. We see that going from minimum transparency in regime U ( $\tau = 0$ ) to maximum transparency ( $\tau = 1$ ) in regime OI, reduces the social loss but increases the central bank loss. We also see that regime OG has the highest loss both for the central bank and for society. Regime S must have the lowest social loss, whereas the central bank loss is then the second highest. The social and central bank rankings of regimes U, OI and S are opposite.

The differences among regimes U, OI and OG are significant and potentially of economic importance. The marginal effect of reputation on expected inflation,  $(k_2 + k_3)$  (cf. (3.21)), is more than twice as high in regime OG than in regime OI. Further, the results imply that

Table 5.2. A numerical example

Regime	$k_0$	$k_2$	$k_3$	$g$	$\text{Var}[z_{t t-1}]$	$E[x_t^2]$			Loss	
						$\pi_t$	$l_t - l_t^*$	$l_t$	$E[L_t]$	$E[L_t^p]$
U	0.56	0.41	0.34	0.37	0.34	1.88	3.15	1.52	2.52	1.70
OI	0.20	0.18	0.18	3.84	0.96	1.45	3.88	1.28	2.66	1.36
OG	1.00	0.50	0.50		0.96	3.46	3.46	1.50	3.46	2.48
S	0.00	0.00	0.00		0.96	1.25	4.21	1.25	2.73	1.25

Note:  $l^* = 1, \beta = 0.95, \rho = 0.7$  and  $\sigma_\varepsilon^2 = \sigma_\theta^2 = \sigma_\eta^2 = 1$ . In regime U,  $\tau = 0$ .

inflation will be 3 percent above the zero inflation target more than 10 percent of the time in regime OG, but less than 1 percent of the time in regime OI.

#### 5.4 Optimal transparency

Proposition 5.1 shows how social and central bank loss vary when transparency is marginally increased. In this subsection, we examine which degrees of transparency are optimal for society and for the central bank, respectively. Since loss does not change monotonically for all levels of  $\tau$ , proposition 5.1 does not imply that there is any value for  $\beta, \rho$  and the  $\sigma^2$ s for which  $\tau = 1$  minimizes loss. In order to examine this numerically, we need to hold each combination of  $\beta, \rho, l^*$ , and the  $\sigma^2$ s fixed, and examine social and central bank loss for different  $\tau$ s (whereas in proposition 5.1 the  $\beta, \rho$ , the  $\sigma^2$ s and the  $\tau$  all vary for each draw). The result of this examination is summarized in:

**Proposition 5.3.** *For the full parameter space:*

- (i) *Full transparency of intention minimizes the social loss for 97.3 percent of the parameter space. The social loss is always minimized at either  $\tau = 1$  or  $\tau = 0$ .*
- (ii) *Full transparency of intention minimizes the central bank loss for 79.5 percent of the parameter space, whereas minimum transparency minimizes it for 18.6 percent. An intermediate degree of transparency is best for central bank loss for 1.9 percent.*
- (iii) *The optimal transparency is always at least as high for society as for the central bank. For 15.9 percent of the parameter space,  $\tau = 0$  minimizes central bank loss while  $\tau = 1$  minimizes social loss.*

This proposition is shown numerically. Thus, full transparency of intention is generally best for society. For the full parameter space, it is generally better for the central bank as well. The reduction in the average inflation bias caused by increased transparency, together with some discounting by the central bank, contributes to this. Throughout the full parameter space, the optimal degree of transparency is at least as high for society as for the central bank.

Social and central bank preferences sharply conflict on a strikingly large portion of the parameter space (15.8 percent): the bank wants minimum transparency and the public wants full transparency. We can shed further light on this phenomenon by considering the “small” parameter space which differs from the full space only in imposing that the bank is patient ( $\beta = 0.99999$ ) and has no average bias ( $l^* = 0$ ). For this case we have

**Proposition 5.4.** *With a patient central bank with no average inflation bias, central bank loss is monotonically increasing with transparency, while social loss is monotonically falling with transparency. Thus,  $\tau = 0$  minimizes central bank loss and maximizes social loss, and  $\tau = 1$  maximizes central bank loss and minimizes social loss.*

This result is clear from Table 5.1.<sup>25</sup> The fact that the public likes transparency is not much different from the full parameter space. The reason why the central bank wants minimum transparency has to do with two features. One benefit of higher transparency to the central bank is a lower average inflation bias. With  $l^* = 0$ , this benefit is absent. In this case, the cost to the bank of increasing transparency, in terms of a more limited ability to generate a correlation between employment and  $z_t$ , comes to dominate. When a shock drives  $z_t$  up,  $z_t$  remains high and because the public learns slowly, the bank can push  $l_t$  up for several periods. For a patient bank, the current and future costs of reducing this ability makes any increase in transparency bad.<sup>26</sup>

We think that this case of low bias and a relatively patient bank may be very relevant in reality. Thus, this suggests a possible conflict between the central bank and society regarding transparency of actions.

Checking which  $\tau$  minimizes the unconditional loss functions is a simple and well-defined exercise, but it raises a great many subtle questions that are beyond the scope of this paper. For example, if we consider a central bank choosing the optimal  $\tau$  once-and-for-all, we are implicitly assuming that the central bank has a commitment technology regarding transparency, but not with regard to the policy rule itself. While CM make a similar assumption in discussing optimal control-error noise, the practical relevance of this assumption is questionable. The alternative of modelling the setting of transparency under discretion takes us far beyond the current framework. CM’s work and other related results raise other questions. For example, if we view the public

---

<sup>25</sup> We suspect, but have not confirmed that for  $\beta < 1$ , there is some tiny portion of the parameter space for which the bank prefers nonzero transparency. This region is smaller than the tolerance for our absolute statements given in section 3.4.

<sup>26</sup> Of course, in the limiting case as  $\beta\rho \rightarrow 1$ , all regimes converge to the social optimum under commitment. Thus, a patient bank with no average bias always disagrees with the public over the optimum transparency. In the limit as the persistence of the goal goes to one, this disagreement becomes moot.

as choosing  $\tau$  and imposing it on the central bank, we must consider Lewis's [21] argument that the central bank could offset increased transparency by increasing the variance of the control error. This suggests a game between the central bank and those regulating it. A related paper in progress, [9], takes up all these issues.

## 6 Conclusions

In this paper we examine the importance of transparency and credibility in monetary policy in a model where the central bank announces a zero-inflation policy. Part of the central bank's goal, in this case an employment target, are idiosyncratic, stochastic and time-varying, and private information to the central bank. The private sector attempts to deduce the central bank's goals from available information and forms rational expectations of future inflation. We demonstrate an equilibrium with time-varying central-bank reputation and time-varying credibility the of zero-inflation announcements. Credibility is defined as negatively related to the distance between the private sector's inflation expectations and the bank's announced inflation target. Our notion of transparency is related to the size of an (to the private sector) unobservable component of an unavoidable control error in monetary policy; increased transparency (smaller unobservable noise) improves the precision of private-sector inference about central bank goals, and makes the bank's reputation and the private sector's inflation expectations more correct.

We believe that we have improved upon the previous literature, Cukierman and Meltzer [7], by acknowledging an explicit stabilization objective for output or employment, and by distinguishing transparency from control-error variance. The former makes central bank policy depend on its reputation and credibility; the latter we believe is necessary for avoiding confusion between transparency and control in monetary policy.

We examine some frequent claims in the literature on practical monetary policy. One claim is that a low-credibility bank, everything else equal, should conduct a more restrictive policy than a high-credibility bank. Specifying "everything else equal" as implying the same true employment target, the same shocks, and the same degree of transparency, we find that, on the contrary, a low-credibility bank (that is, for which private-sector inflation expectations exceed the announced zero inflation rate by more) will induce higher inflation than a high-credibility bank, and hence in this sense conduct a *more* inflationary policy. However, the low-credibility bank at the same time induces lower inflation than expected, and therefore lower employment; hence, in this sense it conducts a *less* expansionary policy than a high-credibility bank.

A second claim is that a low-credibility bank has less flexibility to respond to shocks, in order not to deteriorate its credibility further. In contrast, we find that low and high-credibility banks react in the same way to supply shocks and shocks to the employment target; it is not the case that the low-credibility bank has less scope to stabilize supply shocks, nor does it more urgently build up credibility than a high-credibility bank. This is due to the linearity of the policy rule, and suggests that the second claim must rest on nonlinearities not present in the current model.

A third claim is that increased transparency increases credibility and improves the policy outcome. With regard to this claim, we show that increased transparency makes the central bank's reputation and private-sector inflation expectations more sensitive to the central bank's actions. This generally increases the costs for the bank of deviating from the announced zero-inflation policy, and hence deters the bank further from attempting to fulfill its idiosyncratic employment target. As a result, variability of inflation and employment falls, and any average inflation bias is reduced. These changes generally (but not always) increase social welfare. In many cases, however, increased transparency leads to a worse outcome for the bank; thus, the central bank's preference and social preference for transparency may diverge. Thus, society may prefer to decide on the level of transparency in monetary policy, rather than to delegate this decision to the central bank, since the latter, left on its own, may bring less than socially optimal transparency.<sup>27</sup>

The fact that increased transparency makes the bank's optimal policy closer to the social optimum may throw some light on McCallum's [22] criticism of discretion equilibria in monetary policy. McCallum argues that the problems arising in discretion equilibria will simply not arise in practice because central bankers see the value of the policy consistent with commitment and can *just do it*. If we are to maintain the equilibrium framework, this can only be interpreted as the belief that there is some heretofore unmodelled aspect of preferences or commitment mechanisms that alter the equilibrium outcomes—without some such element *just do it* is not an equilibrium. We are sympathetic to this view: some commitment mechanism may exist. We would like to see this mechanism specified and discussed, however, because we are also sympathetic with Canzoneri's [5] view that in the presence of private information, commitment would be hard to sustain. In our model, such private information exists, but we see our model as providing some insight as to why the problems with discretion might not be as bad as it seems in

---

<sup>27</sup> We recall Milton Friedman's response to Fischer [12], footnote 52, on central bankers' loss functions: "From revealed preference, I suspect that by far and away the two most important variables in their loss functions are avoiding accountability on one hand and achieving public prestige on the other."

some models. In the model, the central bank's concern about its reputation creates an incentive to behave closer to the socially optimal policy. Increased transparency makes this incentive stronger. Indeed, a very patient central bank with very persistent idiosyncratic deviations from the social employment goal would, in the limit, follow the socially optimal policy. Thus, credibility and transparency may matter in ways that make the discretion equilibrium similar to the socially preferred equilibrium, in the absence of an explicit commitment mechanism.<sup>28</sup>

With regard to transparency, we find an especially intriguing result: Consider the case of “extreme” transparency, when the central bank's idiosyncratic goals can be observed directly, rather than be inferred from the central bank's actions. Then the central bank's reputation becomes completely independent from the bank's actions, and the central bank loses an important constraint on its behavior. As a result, a bad equilibrium arises, with high inflation bias and high variability of inflation and employment. Thus, extreme transparency may appear counter-productive. However, extreme transparency can be productive in another sense: If society could observe the central bank's idiosyncratic goals directly, it might not tolerate such idiosyncratic goals, and instead find ways to enforce its own goals on the central bank.

There are some obvious qualifications to our results, some of which may be suitable for future work. In a separate paper, [9], we are taking up a more thorough treatment of the optimal transparency considering the issues of commitment and the incentives of the central bank to cheat. We noted above the importance of assuming that supply shocks are known *ex post* by the private sector, and therefore not confused with shocks to the central bank's goals. Obviously, such confusions, as well as confusions between temporary and persistent shocks, would complicate the private sector's signal extraction problem and possibly modify some of our results. Finally, we have restricted the private sector to believe that the central bank's policy rule is linear; as a consequence, it is optimal for the central bank to restrict itself to a linear rule. Relaxing this assumption, as well as generalizing other aspects beyond our fundamentally linear-quadratic problem, seems beyond what is currently feasible, at least to us.

---

<sup>28</sup> Others interpret McCallum as implicitly relying on trigger strategy equilibria, an interpretation that McCallum rejects.

## A The Kalman filter

The transition equation is (cf. Harvey [16])

$$z_t = \rho z_{t-1} + \theta_t.$$

The measurement equation is

$$y_t = \kappa z_t + \nu_t,$$

where  $\kappa = k_2$  in our baseline regime. The initial estimates are

$$\begin{aligned} \mathbf{E}_{t-1}^p [z_{t-1}] &= z_{t-1} \\ \text{Var}_{t-1}^p [z_{t-1}] &= P_{t-1|t-1}. \end{aligned}$$

The optimal prediction of the mean is

$$z_{t|t-1} = \rho z_{t-1|t-1},$$

and of the variance

$$P_{t|t-1} = \mathbf{E}_{t-1}^p \left[ \left( z_t - z_{t|t-1} \right)^2 \right] = \rho^2 P_{t-1|t-1} + \sigma_\theta^2.$$

The updating equations are

$$\begin{aligned} z_{t|t} &= z_{t|t-1} + \frac{\kappa P_{t|t-1}}{F_t} (y_t - \kappa z_{t|t-1}) \\ F_t &\equiv \kappa^2 P_{t|t-1} + \sigma_\nu^2 \\ P_{t+1|t} &= \rho^2 \left( P_{t|t-1} - \frac{\kappa^2 P_{t|t-1}^2}{F_t} \right) + \sigma_\theta^2. \end{aligned}$$

In a steady state, that is, when the optimal prediction of the variance has converged, we have  $P_{t+1|t} = P_{t|t-1} = P$ , which implies

$$P = \rho^2 \left( P - \frac{\kappa^2 P^2}{\kappa^2 P + \sigma_\nu^2} \right) + \sigma_\theta^2 = \rho^2 \frac{\sigma_\nu^2 P}{\kappa^2 P + \sigma_\nu^2} + \sigma_\theta^2.$$

This is a Riccati equation for  $P$ , which has one positive and one negative root. The positive root is

$$P = P(\kappa) \equiv \sqrt{\left( \frac{(1 - \rho^2) \frac{\sigma_\nu^2}{\kappa^2} - \sigma_\theta^2}{2} \right)^2 + \sigma_\theta^2 \frac{\sigma_\nu^2}{\kappa^2}} - \frac{(1 - \rho^2) \frac{\sigma_\nu^2}{\kappa^2} - \sigma_\theta^2}{2} > 0. \quad (\text{A.1})$$

The updating equation can then be written

$$z_{t+1|t} = \rho z_{t|t-1} + g \left[ \kappa (z_t - z_{t|t-1}) + \nu_t \right] \quad (\text{A.2})$$

$$= (\rho - g\kappa) z_{t|t-1} + g (\kappa z_t + \nu_t), \quad (\text{A.3})$$

where  $g$ , the Kalman gain, is given by

$$g = g(\kappa) \equiv \rho \frac{\kappa P(\kappa)}{\kappa^2 P(\kappa) + \sigma_\nu^2}. \quad (\text{A.4})$$



If  $\rho$  and  $\kappa$  are both positive,

$$0 \leq g(\kappa)\kappa \equiv \rho \frac{\kappa^2 P(\kappa)}{\kappa^2 P(\kappa) + \sigma_\nu^2} \leq \rho. \quad (\text{A.5})$$

The relation between  $P$  and  $g\kappa$  can be expressed in several different ways, for instance,

$$\begin{aligned} P &= \frac{\sigma_\theta^2 + g^2 \sigma_\nu^2}{1 - (\rho - g\kappa)^2} = \sigma_\theta^2 + g^2 \sigma_\nu^2 + (\rho - g\kappa)^2 P \\ &= \sigma_\theta^2 + \rho^2 \frac{\sigma_\nu^2 P}{\kappa^2 P + \sigma_\nu^2} = \frac{g\kappa \sigma_\nu^2}{\kappa^2 (\rho - g\kappa)} = \frac{\sigma_\theta^2}{1 - \rho(\rho - g\kappa)}. \end{aligned}$$

### A.1 Dynamics, moments and the signal-to-noise ratio

Introduce the forecast error

$$e_{t+1} \equiv z_{t+1} - z_{t+1|t}$$

and note that

$$\begin{aligned} z_t &= \rho z_{t-1} + \theta_t \\ &= z_{t|t-1} + e_t, \end{aligned}$$

where the components are conditionally and unconditionally orthogonal.

The conditional first and second moments are

$$\begin{aligned} \mathbf{E}_{t-1}^p e_t &= 0 \\ \mathbf{E}_{t-1}^p z_t &= z_{t|t-1} \\ \mathbf{E}_{t-1} e_t &= \rho z_{t-1} - z_{t|t-1} \\ \mathbf{E}_{t-1} z_t &= \rho z_{t-1} \\ \text{Var}_{t-1}^p e_t &= \text{Var}_{t-1}^p z_t = P \\ \text{Var}_{t-1}^p z_{t+1|t} &= g^2 (\kappa^2 \text{Var}_{t-1}^p e_t + \sigma_\nu^2) = g^2 (\kappa^2 P + \sigma_\nu^2) \\ \text{Cov}_{t-1}^p [z_{t|t-1}, e_t] &= \text{Cov}_{t-1}^p [z_{t|t-1}, z_t] = \text{Cov}_{t-1} [z_{t|t-1}, e_t] = \text{Cov}_{t-1} [z_{t|t-1}, z_t] = 0 \\ \text{Var}_{t-1} e_t &= \text{Var}_{t-1} z_t = \sigma_\theta^2 < P \end{aligned}$$

The unconditional moments—identical for the private sector and central bank—are,

$$\begin{aligned} \text{Var} [z_t] &= \frac{\sigma_\theta^2}{1 - \rho^2} \\ \text{Var} [e_t] &= P \\ \text{Var} [z_{t|t-1}] &= \frac{\sigma_\theta^2}{1 - \rho^2} - P \\ \text{Cov} [z_{t|t-1}, z_t] &= \text{Var} [z_{t|t-1}] \\ \text{Cov} [z_{t|t-1}, e_t] &= 0. \end{aligned}$$

In (A.3), call  $y_t = \kappa z_t + \nu_t$  the signal, and  $\nu_t$  the noise. Then the signal-to-noise (variance) ratio,  $s(\kappa)$ , is defined as

$$s(\kappa) = \frac{\text{Var}_{t-1}^p [y_t]}{\text{Var}_{t-1}^p [\nu_t]} = \frac{\kappa^2 P(\kappa) + \sigma_\nu^2}{\sigma_\nu^2} \geq 1.$$

With some algebra,  $P$  and  $g\kappa$  can be expressed in terms of  $s$  as

$$P = (s-1) \frac{\sigma_\nu^2}{\kappa^2} = \frac{s}{s-\rho^2} \sigma_\theta^2 \quad (\text{A.6})$$

$$g\kappa = \rho \frac{\kappa^2 P}{\kappa^2 P + \sigma_\nu^2} = \frac{\rho(s-1)}{s} \quad (\text{A.7})$$

$$\rho - g\kappa = \frac{\rho}{s}. \quad (\text{A.8})$$

It follows that  $P$ ,  $g$  are increasing in  $s$  for given  $\kappa$ , and that  $P$  and  $g\kappa$  are also increasing in  $s$  when  $\kappa$  is endogenous, as in regime U. We note from (A.6) that  $s$  can be expressed without  $\kappa$  as

$$s = \frac{\rho^2 P}{P - \sigma_\theta^2}. \quad (\text{A.9})$$

The dynamics of  $z_{t|t-1}$  can be written in terms of the signal-to-noise ratio as

$$z_{t+1|t} = \rho \left[ \frac{1}{s} z_{t|t-1} + \left(1 - \frac{1}{s}\right) z_t \right] + g\nu_t.$$

## A.2 Dependence of $g\kappa$ on $\kappa$ and $\sigma_\nu^2$

We can show that  $g(\kappa)\kappa$  is increasing in  $\kappa$ , despite the fact that  $P(\kappa)$  is not monotonic in  $\kappa$ . This is because  $\kappa^2$  dominates in the term  $\kappa^2 P(\kappa)$  in (A.5). In order to show this, we note that

$$\begin{aligned} \kappa^2 P(\kappa) &= \sqrt{\left(\frac{(1-\rho^2)\sigma_\nu^2 - \kappa^2\sigma_\theta^2}{2}\right)^2 + \kappa^2\sigma_\theta^2\sigma_\nu^2} - \frac{(1-\rho^2)\sigma_\nu^2 - \kappa^2\sigma_\theta^2}{2} \\ &= \frac{\sqrt{[(1-\rho^2)\sigma_\nu^2 - \kappa^2\sigma_\theta^2]^2 + 4\kappa^2\sigma_\theta^2\sigma_\nu^2} - [(1-\rho^2)\sigma_\nu^2 - \kappa^2\sigma_\theta^2]}{2}. \end{aligned}$$

We can set  $\kappa \equiv 1$  and examine  $\frac{\partial P(1)}{\partial \sigma_\theta^2}$ . We get

$$\begin{aligned} 2P(1) &= \sqrt{[(1-\rho^2)\sigma_\nu^2 - \sigma_\theta^2]^2 + 4\sigma_\theta^2\sigma_\nu^2} - [(1-\rho^2)\sigma_\nu^2 - \sigma_\theta^2] \\ 2\frac{\partial P(1)}{\partial \sigma_\theta^2} &= \frac{1}{2} \frac{1}{\sqrt{\cdot}} \left\{ 2[(1-\rho^2)\sigma_\nu^2 - \sigma_\theta^2](-1) + 4\sigma_\nu^2 \right\} + 1 \\ &= \frac{1}{\sqrt{\cdot}} \left\{ 2\sigma_\nu^2 - [(1-\rho^2)\sigma_\nu^2 - \sigma_\theta^2] \right\} + 1 \\ &= \frac{\left\{ \sqrt{[(1-\rho^2)\sigma_\nu^2 - \sigma_\theta^2]^2 + 4\sigma_\theta^2\sigma_\nu^2} - [(1-\rho^2)\sigma_\nu^2 - \sigma_\theta^2] \right\} + 2\sigma_\nu^2}{\sqrt{\cdot}} > 0, \end{aligned}$$

where we observe that the term within curly brackets is positive. Hence  $\frac{\partial \kappa^2 P(\kappa)}{\partial \kappa} > 0$ .

In order to show  $\partial g\kappa / \partial \sigma_\nu^2 > 0$ , let us start with  $s$ , which we can write

$$s \equiv \frac{\kappa^2 P(\kappa) + \sigma_\nu^2}{\sigma_\nu^2} = \frac{\sqrt{[(1-\rho^2) - \kappa^2\sigma_\theta^2/\sigma_\nu^2]^2 + 4\kappa^2\sigma_\theta^2/\sigma_\nu^2} - [(1-\rho^2) - \kappa^2\sigma_\theta^2/\sigma_\nu^2]}{2} + 1.$$

We realize that

$$\frac{\partial s}{\partial \sigma_\nu^2} = \frac{\partial s}{\partial (\kappa^2 \sigma_\theta^2 / \sigma_\nu^2)} \left( - \frac{\kappa^2 \sigma_\theta^2}{\sigma_\nu^4} \right).$$

However,  $\frac{\partial s}{\partial (\kappa^2 \sigma_\theta^2 / \sigma_\nu^2)} > 0$  by the same argument as for  $\frac{\partial P(1)}{\partial \sigma_\theta^2}$  above. Hence,  $\frac{\partial s}{\partial \sigma_\nu^2} < 0$ . Thus, an increase in the noise variance  $\sigma_\nu^2$  reduces the signal-to-noise ratio  $s$ . By (A.6) and (A.7) this increases  $P$  and  $g\kappa$ . This relation holds generally, and thus, at the optimum when  $\kappa$  is endogenous, as is the case in regime U.

## B Unobservable goal

Using

$$\mathbf{E}_{t-} l_t = i_t - \pi_{t|t-1} + \varepsilon_t,$$

we can write the first-order condition (3.7) as

$$i_t = \frac{1}{2} \left\{ l^* + z_t + \pi_{t|t-1} - \beta \mathbf{E}_{t-} \left[ (\delta_1 + \delta_2 z_{t+1|t} + \delta_5 z_t) \frac{\partial z_{t+1|t}}{\partial i_t} \right] \right\} - \frac{1}{2} \varepsilon_t. \quad (\text{B.1})$$

From (2.12), (3.2) and (3.3) and we have,

$$\pi_{t|t-1} = k_0 + (k_2 + k_3) z_{t|t-1} \quad (\text{B.2})$$

$$\frac{\partial z_{t+1|t}}{\partial i_t} = g \quad (\text{B.3})$$

$$\mathbf{E}_{t-} z_{t+1|t} = (\rho - gk_2) z_{t|t-1} + gk_2 z_t. \quad (\text{B.4})$$

Substituting and collecting terms gives:

$$\begin{aligned} i_t &= \frac{1}{2} (l^* + k_0 - \beta g \delta_1) - \frac{1}{2} \varepsilon_t + \frac{1}{2} [1 - \beta g (gk_2 \delta_2 + \delta_5)] z_t \\ &\quad + \frac{1}{2} [k_2 + k_3 - \beta g (\rho - gk_2) \delta_2] z_{t|t-1} \end{aligned}$$

This function is of the form (2.12) with (3.8)-(3.11).

Now return to the value function using (2.12) and (2.18),

$$\begin{aligned} V(z_{t|t-1}, z_{t-1}) &= \mathbf{E}_{t-1} \left\{ \frac{1}{2} \left[ (k_0 + k_1 \varepsilon_t + k_2 z_t + k_3 z_{t|t-1} + \eta_t)^2 \right. \right. \\ &\quad \left. \left. + ((k_1 + 1) \varepsilon_t + k_2 (z_t - z_{t|t-1}) + \eta_t - l^* - z_t)^2 \right] + \beta V(z_{t+1|t}, z_t) \right\}. \end{aligned}$$

Expansion of the value function gives:

$$\begin{aligned} V(z_{t|t-1}, z_{t-1}) &= \mathbf{E}_{t-1} \left\{ \frac{1}{2} \left[ k_0^2 + k_1^2 \varepsilon_t^2 + k_2^2 z_t^2 + k_3^2 z_{t|t-1}^2 + \eta_t^2 \right] \right. \\ &\quad + k_0 k_2 z_t + k_0 k_3 z_{t|t-1} + k_2 k_3 z_{t|t-1} z_t \\ &\quad + \frac{1}{2} \left[ (k_1 + 1)^2 \varepsilon_t^2 + (k_2 - 1)^2 z_t^2 + k_2^2 z_{t|t-1}^2 + \eta_t^2 + l^{*2} \right] \\ &\quad - (k_2 - 1) k_2 z_{t|t-1} z_t - (k_2 - 1) l^* z_t + k_2 l^* z_{t|t-1} \\ &\quad \left. + \beta \delta_0 + \beta \delta_1 z_{t+1|t} + \frac{1}{2} \beta \delta_2 z_{t+1|t}^2 + \beta \delta_3 z_t + \frac{1}{2} \beta \delta_4 z_t^2 + \beta \delta_5 z_{t+1|t} z_t \right\}. \end{aligned}$$

The following expressions are useful in evaluating this expectation:

$$\mathbf{E}_{t-1} z_{t+1|t} = (\rho - gk_2) z_{t|t-1} + \rho g k_2 z_{t-1} \quad (\text{B.5})$$

$$\mathbf{E}_{t-1} z_{t+1|t}^2 = (\rho - gk_2)^2 z_{t|t-1}^2 + (gk_2)^2 (\rho^2 z_{t-1}^2 + \sigma_\theta^2) + 2\rho(\rho - gk_2) g k_2 z_{t|t-1} z_{t-1} + g^2 \sigma_\nu^2 \quad (\text{B.6})$$

$$\begin{aligned} \mathbf{E}_{t-1} z_{t+1|t} z_t &= \mathbf{E}_{t-1} \left[ (\rho - gk_2) z_{t|t-1} + g(k_2 z_t + \eta_t) \right] z_t \\ &= \rho(\rho - gk_2) z_{t|t-1} z_{t-1} + gk_2 (\rho^2 z_{t-1}^2 + \sigma_\theta^2). \end{aligned} \quad (\text{B.7})$$

Thus, evaluating the expectations yields

$$\begin{aligned} V(z_{t|t-1}, z_{t-1}) &= \frac{1}{2} \left[ k_0^2 + \frac{1}{4} \sigma_\epsilon^2 + k_2^2 (\rho^2 z_{t-1}^2 + \sigma_\theta^2) + k_3^2 z_{t|t-1}^2 + \sigma_\eta^2 \right] \\ &\quad + k_0 k_2 \rho z_{t-1} + k_0 k_3 z_{t|t-1} + k_2 k_3 \rho z_{t|t-1} z_{t-1} + \\ &\quad \frac{1}{2} \left[ \frac{1}{4} \sigma_\epsilon^2 + (k_2^2 - 2k_2 + 1)(\rho^2 z_{t-1}^2 + \sigma_\theta^2) + k_2^2 z_{t|t-1}^2 + \sigma_\eta^2 + l^{*2} \right] \\ &\quad - (k_2 - 1) k_2 \rho z_{t|t-1} z_{t-1} - (k_2 - 1) l^* \rho z_{t-1} + k_2 l^* z_{t|t-1} \\ &\quad + \beta \delta_0 + \beta \delta_1 (\rho - gk_2) z_{t|t-1} + \beta \delta_1 \rho g k_2 z_{t-1} \\ &\quad + \frac{1}{2} \beta \delta_2 \left[ (\rho - gk_2)^2 z_{t|t-1}^2 + \rho^2 (gk_2)^2 z_{t-1}^2 + (gk_2)^2 \sigma_\theta^2 + g^2 \sigma_\nu^2 \right] \\ &\quad + \beta \delta_2 \rho (\rho - gk_2) g k_2 z_{t|t-1} z_{t-1} + \beta \delta_3 \rho z_{t-1} + \frac{1}{2} \beta \delta_4 (\rho^2 z_{t-1}^2 + \sigma_\theta^2) \\ &\quad + \beta \delta_5 \left[ ((\rho - gk_2) z_{t|t-1} + \rho g k_2 z_{t-1}) \rho z_{t-1} + g k_2 \sigma_\theta^2 \right]. \end{aligned}$$

We can obtain expressions for the  $\delta$ s by collecting the relevant terms. The constant:

$$\begin{aligned} \delta_0 &= \frac{1}{2} \left[ k_0^2 + \frac{1}{4} \sigma_\epsilon^2 + k_2^2 \sigma_\theta^2 + \sigma_\eta^2 \right] + \frac{1}{2} \left[ \frac{1}{4} \sigma_\epsilon^2 + (k_2^2 - 2k_2 + 1) \sigma_\theta^2 + \sigma_\eta^2 + l^{*2} \right] \\ &\quad + \beta \delta_0 + \frac{1}{2} \beta \delta_2 \left[ (gk_2)^2 \sigma_\theta^2 + g^2 \sigma_\nu^2 \right] + \frac{1}{2} \beta \delta_4 \sigma_\theta^2 \\ \delta_0 &= \frac{\frac{1}{2} l^{*2} + k_0^2 + \frac{1}{2} \sigma_\epsilon^2 + (2k_2^2 - 2k_2 + 1 + \beta (gk_2)^2 \delta_2 + \beta \delta_4) \sigma_\theta^2}{1 - \beta} \\ &\quad + \frac{1}{2} \frac{[2 + \beta(1 - \tau)g^2 \delta_2 + 2\beta g k_2 \delta_5] \sigma_\eta^2}{1 - \beta}, \end{aligned} \quad (\text{B.8})$$

where we have used (2.10). Collect terms in  $z_{t|t-1}$ :

$$\begin{aligned} \delta_1 &= k_0 k_3 + l^* k_2 + \beta(\rho - gk_2) \delta_1 = (l^* - \beta g \delta_1) k_3 + l^* k_2 + \beta(\rho - gk_2) \delta_1 \\ &= l^* (k_2 + k_3) - \beta g k_3 \delta_1 + \beta(\rho - gk_2) \delta_1 \end{aligned}$$

where we have used (3.8). Thus

$$\delta_1 = \frac{l^* (k_2 + k_3)}{1 - \beta(\rho - gk_2) + \beta g k_3}. \quad (\text{B.9})$$

Collect terms in  $z_{t|t-1}^2$ :

$$\delta_2 = k_2^2 + k_3^2 + \beta(\rho - gk_2)^2 \delta_2$$

$$\delta_2 = \frac{k_2^2 + k_3^2}{1 - \beta(\rho - gk_2)^2}. \quad (\text{B.10})$$

Collect terms in  $z_{t-1}$ :

$$\begin{aligned} \delta_3 &= k_0 k_2 \rho + l^* \rho (1 - k_2) + \beta \rho g k_2 \delta_1 + \beta \rho \delta_3 \\ \delta_3 &= \frac{\rho k_0 k_2 + \rho l^* (1 - k_2) + \beta \rho g k_2 \delta_1}{1 - \beta \rho} = \frac{\rho (l^* - \beta g \delta_1) k_2 + \rho l^* (1 - k_2) + \beta \rho g k_2 \delta_1}{1 - \beta \rho} \\ &= \frac{\rho l^*}{1 - \beta \rho}, \end{aligned} \quad (\text{B.11})$$

where we have used (3.8). Collect terms in  $z_{t-1}^2$ :

$$\begin{aligned} \delta_4 &= k_2^2 \rho^2 + (k_2^2 - 2k_2 + 1) \rho^2 + \beta \rho^2 (gk_2)^2 \delta_2 + \beta \rho^2 \delta_4 + 2\beta \rho^2 g k_2 \delta_5 \\ \delta_4 &= \rho^2 \frac{1 - 2k_2 + 2k_2^2 + \beta (gk_2)^2 \delta_2 + 2\beta g k_2 \delta_5}{1 - \beta \rho^2}. \end{aligned} \quad (\text{B.12})$$

Collect terms in  $z_{t|t-1} z_{t-1}$ :

$$\begin{aligned} \delta_5 &= \rho k_2 k_3 - (k_2 - 1) k_2 \rho + \beta \rho (\rho - gk_2) g k_2 \delta_2 + \beta \rho (\rho - gk_2) \delta_5 \\ \delta_5 &= \frac{\rho k_2}{1 - \beta \rho (\rho - gk_2)}, \end{aligned} \quad (\text{B.13})$$

where we have used (3.11).

## B.1 Existence and uniqueness

We now have a simultaneous set of 9 equations, 3 for the  $k$ s and 6 for the  $\delta$ s ( $k_1 = -\frac{1}{2}$  is known). We first show that we can rewrite that system as a single equation for  $k_2$  in terms of itself,

$$k_2 = f(k_2),$$

and equations giving the eight remaining  $k$ s and  $\delta$ s in terms of  $k_2$ .

First, we get an expression for  $k_3$  in terms of  $k_2$  only. Since  $g$  by (A.4) depends only on  $k_2$ , and  $\delta_2$  by (B.10) depends only on  $k_2$  and  $k_3$ , (3.11) can be written as an expression in  $k_2$  and  $k_3$  only. Taking  $k_2$  as fixed, the equation is a quadratic in  $k_3$ :

$$k_3 = k_2 - \frac{\beta g(k_2) [\rho - g(k_2) k_2]}{1 - \beta (\rho - g(k_2) k_2)^2} k_2^2 - \frac{\beta g(k_2) [\rho - g(k_2) k_2]}{1 - \beta (\rho - g(k_2) k_2)^2} k_3^2$$

or

$$\begin{aligned} 0 &= \frac{\beta g(k_2) k_2 [\rho - g(k_2) k_2]}{[1 - \beta (\rho - g(k_2) k_2)^2] k_2} k_3^2 + k_3 - k_2 \left[ 1 - \frac{\beta g(k_2) k_2 [\rho - g(k_2) k_2]}{[1 - \beta (\rho - g(k_2) k_2)^2]} \right] \\ &= \frac{A(k_2)}{k_2} k_3^2 + k_3 - k_2 [1 - A(k_2)] \end{aligned}$$

which has solutions of the form,

$$k_3 = \frac{-1 \pm \sqrt{1 + 4A(k_2)[1 - A(k_2)]}}{2A(k_2)} k_2$$

where

$$A(k_2) = \frac{\beta g(k_2) k_2 [\rho - g(k_2) k_2]}{1 - \beta [\rho - g(k_2) k_2]^2}.$$

We note that we can write  $A(k_2)$  as

$$A(k_2) = \frac{\beta [\rho - (\rho - g(k_2) k_2)] [\rho - g(k_2) k_2]}{1 - \beta [\rho - g(k_2) k_2]^2} = \frac{\beta \rho [\rho - g(k_2) k_2] - \beta [\rho - g(k_2) k_2]^2}{1 - \beta [\rho - g(k_2) k_2]^2},$$

hence

$$0 \leq A(k_2) < 1, \quad (\text{B.14})$$

since  $0 \leq \rho - g(k_2) k_2 < 1$ .

We choose the positive root for  $k_3$ . We have not proved, in general, that the negative root does not give an alternative equilibrium of the model, but we have two arguments for ignoring this root. First, McCallum [23] argues that we should consider solutions for which the coefficients of the policy rule are continuous in the parameters of the problem. This rules out the negative root. For  $\sigma_v^2 = 0$ , we know that  $k_3 = k_2$ . Imposing continuity of  $k_3$  in  $\sigma_v^2$  at  $\sigma_v^2 = 0$ . For  $\sigma_v^2 \rightarrow 0$ ,  $g(k_2) k_2 \rightarrow \rho$  and  $A(k_2) \rightarrow 0$ . It follows that the root above must be

$$k_3(k_2) = \frac{\sqrt{1 + 4A(k_2)[1 - A(k_2)]} - 1}{2A(k_2)} k_2, \quad (\text{B.15})$$

since the other root does not have a limit for  $\sigma_v^2 \rightarrow 0$ . Furthermore, we note that under (B.14)

$$0 \leq \frac{\sqrt{1 + 4A(k_2)[1 - A(k_2)]} - 1}{2A(k_2)} < 1. \quad (\text{B.16})$$

The second argument rests on the fact that for particular parameter values, one can rule out the negative root by showing that a one-period deviation from the implied policy rule decreases the central bank's loss. Using the approach described in Appendix E, we verified numerically that the negative root is not an equilibrium.

Thus, we have  $k_3$  in terms of  $k_2$  alone, and substituting this expression for  $k_3$  into the formulae for  $\delta_2$  gives an expression for  $\delta_2$  in terms of  $k_2$  alone. By (B.13)  $\delta_5$  depends only on  $k_2$ . Recursive substitution using these results gives expressions for the other  $\delta$ s and  $k_0$ . Finally, substituting the expressions for  $\delta_2$  and  $\delta_5$  into (3.10) gives the desired equation for  $k_2$ :

$$k_2 = f(k_2) \equiv \frac{\frac{1 - \beta \rho^2}{1 - \beta \rho^2 + \beta \rho g(k_2) k_2}}{2 + \beta g(k_2)^2 \delta_2(k_2)}. \quad (\text{B.17})$$

Now existence and uniqueness are only a question of whether there are zero, one, or more solutions to (B.17). Since  $0 < \frac{1 - \beta \rho^2}{1 - \beta \rho^2 + \beta \rho g(k_2) k_2} < 1$  and  $\beta g(k_2)^2 \delta_2(k_2) > 0$ , we have

$$0 < f(k_2) < \frac{1}{2}$$

for all  $k_2$ . Thus, any solution to (B.17) must be in  $[0, \frac{1}{2}]$ . Further, since  $f$  is continuous for  $k_2 \in [0, \frac{1}{2}]$ , (B.17) must have at least one solution. Arguments about uniqueness are discussed in the text and Appendix E.

## B.2 Proof that $k_2$ and $k_3$ go to zero when $\beta\rho$ goes to one

First, we show that  $\lim_{\beta\rho \rightarrow 1} k_2 = 0$ . Denote the numerator in (B.17) by  $N$ , so that

$$\frac{1}{N} = \frac{1 - \beta\rho^2 + \beta\rho g(k_2)k_2}{1 - \beta\rho^2} = 1 + \frac{\beta\rho g(k_2)k_2}{1 - \beta\rho^2},$$

or, using  $g(k_2)k_2 = \frac{\rho k_2^2 P(k_2)}{k_2^2 P(k_2) + \sigma_\nu^2}$ ,

$$\frac{1}{N} = 1 + \frac{\beta\rho^2 k_2^2 P(k_2)}{(1 - \beta\rho^2)(k_2^2 P(k_2) + \sigma_\nu^2)}.$$

Now let  $\beta\rho \rightarrow 1$ , in which case  $\beta \rightarrow 1$  and  $\rho \rightarrow 1$  (since they are both bounded above by one), and  $\beta\rho^2 \rightarrow 1$ . Assume, contrary to the desired result, that  $k_2$  is bounded away from zero as  $\beta\rho^2 \rightarrow 1$ . Since  $P(k_2)$  is bounded below by  $\sigma_\nu^2$ , we have  $\frac{1}{N} \rightarrow \infty$  implying  $N \rightarrow 0$ . Since the denominator of (B.17) is bounded below by 2,  $N \rightarrow 0$  implies that  $k_2 \rightarrow 0$ , a contradiction.

Since, by (B.15) and (B.16),  $0 \leq k_3 < k_2$ , it follows directly that  $\lim_{\beta\rho \rightarrow 1} k_3 \rightarrow 0$ .

## C Regime OI

From (A.4) and (A.1) it is clear that  $gk_2 = \rho$  when  $\sigma_\nu^2 = 0$ . Using this fact and (3.11), we have  $k_2 = k_3$ . Substituting in (B.9), (B.10), and (B.13), gives

$$\delta_1 = \frac{2l^*k_2}{1 + \beta\rho} \tag{C.1}$$

$$\delta_2 = 2k_2^2 \tag{C.2}$$

$$\delta_5 = \rho k_2. \tag{C.3}$$

Now from (3.8) and (C.1), using  $gk_2 = \rho$ ,

$$k_0 = l^* - \frac{2\beta\rho l^*}{1 + \beta\rho} = l^* \frac{1 - \beta\rho}{1 + \beta\rho}.$$

From (3.9),  $k_1 = -\frac{1}{2}$ , and from (3.10), (C.2) and (C.3), using  $gk_2 = \rho$ ,

$$k_2 = \frac{1 - \beta g \rho k_2}{2 + \beta g^2 (2k_2^2)} = \frac{1 - \beta\rho^2}{2(1 + \beta\rho^2)}.$$

These together imply the results in the text.

## D The CM regime

Our baseline regime is closely related to the model of Cukierman and Meltzer [7], Cukierman [6] (Part II, especially Chapters 8–10), and Levine and Pearlman [20]. The information structure in the CM regime is the same as in the U regime above. The only differences are that  $\tau = 0$  and the CM period loss function is

$$L_t^{\text{CM}} = \frac{1}{2} i_t^2 - (A + z_t) (\pi_t - \pi_{t-1}). \tag{D.1}$$

With  $\tau$  and  $\sigma_\eta^2$  fixed, this loss function will give rise to the same first-order condition (after taking expectations  $\mathbf{E}_{t-}$ ) as

$$\begin{aligned} L_t^{\text{CM}'} &= \frac{1}{2}(i_t + \eta_t)^2 - (A + z_t) (\pi_t - \pi_{t|t-1} + \varepsilon_t) \\ &= \frac{1}{2}\pi_t^2 - (A + z_t) l_t. \end{aligned} \quad (\text{D.2})$$

Thus, since the CM loss function is linear in unanticipated inflation, it is implicitly linear in employment: there is no objective to *stabilize* employment.

When  $L_t$  is given by (D.2), the first-order condition for (3.6) is

$$i_t - A - z_t + \beta \mathbf{E}_{t-} \left[ (\delta_1 + \delta_2 z_{t+1|t} + \delta_5 z_t) \frac{\partial z_{t+1|t}}{\partial i_t} \right] = 0, \quad (\text{D.3})$$

where the  $\delta$ s are the coefficients in the quadratic value function of the form (3.5). From (D.3), (B.3) and (B.4) we get

$$i_t = A - \beta g \delta_1 + [1 - \beta g (\delta_2 g k_2 + \delta_5)] z_t - \beta g (\rho - g k_2) \delta_2 z_{t|t-1}.$$

This is a policy rule on the form (2.12), if the coefficients fulfill

$$k_0 = A - \beta g \delta_1 \quad (\text{D.4})$$

$$k_1 = 0 \quad (\text{D.5})$$

$$k_2 = \frac{1 - \beta g \delta_5}{1 + \beta g^2 \delta_2} \quad (\text{D.6})$$

$$k_3 = -\beta g (\rho - g k_2) \delta_2. \quad (\text{D.7})$$

Substitution of (2.12) and (D.4)–(D.7) into (3.4) gives, with the same method as in appendix B,

$$\delta_2 = k_3 = 0 \quad (\text{D.8})$$

$$\delta_1 = \frac{A k_2}{1 - \beta (\rho - g k_2)} \quad (\text{D.9})$$

$$\delta_5 = \frac{\rho k_2}{1 - \beta \rho (\rho - g k_2)}. \quad (\text{D.10})$$

Thus,

$$k_0 = \frac{1 - \beta \rho}{1 - \beta (\rho - g k_2)} A \quad (\text{D.11})$$

$$k_1 = k_3 = 0 \quad (\text{D.12})$$

$$k_2 = \frac{1 - \beta \rho}{1 - \beta (\rho - g k_2)}. \quad (\text{D.13})$$

The most important difference from regime U is that in CM,  $k_1 = k_3 = 0$ . Obviously,  $k_1 = 0$  because the central bank does not care about the effect of the supply shock on employment. Essentially the same reason accounts for the central bank disregarding its reputation ( $k_3 = 0$ ) in forming policy: Consider regime U in a situation when  $l^* = 0$ ,  $z_t = 0$ , and suppose that  $k_3 = 0$  and  $\varepsilon_t = 0$ . The period loss function with  $z_t = l^* = l_t^* = 0$  is

$$\frac{1}{2}(\pi_t^2 + l_t^2).$$



Under these assumptions,  $i_t = 0$ , and  $E\pi_t^2$  is at its minimum of  $\sigma_\eta^2$ . The central bank could lower the period loss by a one-period deviation moving  $i_t$  in the same direction as the sign of  $z_{t|t-1}$ . Take  $i_t = k_3 z_{t|t-1}$  for some small positive  $k_3$ . Lowering  $i_t$  from zero raises  $E\pi_t^2$  but lowers  $El_t^2$  by more. By deviating from the supposed rule and moving  $i_t$  in the direction of what the public expects, the gain from smoothing employment outweighs the rise in squared inflation. This one-period deviation will have a small cost in terms of slowing public learning ( $E_{t-}\partial z_{t+1|t}/\partial i_t > 0$ ), but this term is outweighed. In the CM case, with  $A = z_t = 0$ , the period loss function is simply  $\frac{1}{2}\pi_t^2$ . No employment-stabilizing benefit comes from moving  $i_t$  in the direction the public expects, and the bank has no incentive to use any  $k_3$  other than  $k_3 = 0$ .

## E The numerical work

For a given value of the parameters  $(\beta, \rho, \tau, \sigma_\eta^2, \sigma_\theta^2, l^*)$ , we solve regime U by searching over  $k_2 \in [0, \frac{1}{2}]$  for a  $k_2$  satisfying  $k_2 = f(k_2)$ . Regimes OI and OG can be solved by direct computation. After solving the models, any aspect of the models for which we have formulae can be computed directly. This includes all the results about the values of the  $k$ s, the loss functions,  $g$ , and  $P$ . In particular, the derivatives are all evaluated with analytic formulae.

Three items for which we state numerical results cannot be computed directly: (1) Uniqueness of the solution to  $k_2 = f(k_2)$ , and (2) The incentive to deviate from the possible equilibrium with the negative  $k_3$  root, 3) Verifying what  $\tau$  is optimal for any value for the other parameters.

Numerical uniqueness is checked by computing  $f(k_2)$  for 100 evenly spaced points in the interval  $[0.0001, 0.5]$  and checking whether  $f$  is monotonically declining over the range for those points.

To test the incentive to deviate under the negative  $k_3$  root, we first solve the model for the  $k$ s,  $\delta$ s, and  $g$  taking the negative root for  $k_3$ . We then repeat the following steps for a wide range of the state variables  $z_t$  and  $z_{t|t-1}$  (these are the state variables as of time  $t-$ , which is relevant in what follows): (i) Evaluate the central bank loss under the implied policy rule seen from time  $t-$ , after  $\theta$  and  $\varepsilon = 0$  are drawn at  $t$ , but before  $\eta$  is drawn. (ii) Evaluate the central bank loss seen from time  $t-$  (with  $\varepsilon = 0$ ) from setting  $i_t$  equal to various arbitrarily chosen values, but returning to the policy rule from  $t+1$  onward. If for any  $(z_t, z_{t|t-1})$  pair, there is an  $i_t$  that dominates the policy rule, we have proved that the negative  $k_3$  root is not an equilibrium for this parameter value. In 100,000 draws, about 3 percent of draws would not solve at all with the negative  $k_3$  root; for all the remaining draws, the negative root did not constitute an equilibrium.

For the optimal  $\tau$ , we draw a value for the other parameters and check the value of the two loss functions at 100 evenly spaced points between zero and one. The smallest loss is taken as the optimum.

Four separate numerical experiments were performed: There was one run for the full parameter space and one for the small parameter space in which all aspects except the optimal  $\tau$  and the validity of the negative  $k_3$  root were checked. There was one run checking the optimal  $\tau$  for the full parameter space (since the derivatives of loss with respect to  $\tau$  for the small parameter space were of one sign, the optimal  $\tau$  results follow without further computation). Finally, there was one run checking the validity of the negative  $k_3$  root. In each case, for a small number of draws, numerical instability for certain extreme parameter values kept us from solving the model at all. For the four experiments, this problem arose for 63, 66, 152, and 3,319 draws out of 100,000, respectively.

## References

- [1] Backus, David, and John Driffill (1985), "Inflation and Reputation," *American Economic Review* 75, 530-538.
- [2] Ball, Laurence (1994), "What determines the sacrifice ratio," in N. Gregory Mankiw, ed., *Monetary Policy*, 155-194.
- [3] Barro, Robert, and David Gordon (1983), "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy* 91, 589-610.
- [4] Bernanke, Ben, and Frederic Mishkin (1992), "Central Bank Behavior and the Strategy of Monetary Policy: Observations from Six Industrialized Countries," *NBER Macroeconomics Annual* 7, 183-228.
- [5] Canzoneri, Matthew B. (1985), "Monetary Policy Games and the Role of Private Information," *American Economic Review* 75, 1056-1070.
- [6] Cukierman, Alex (1992), *Central Bank Strategy, Credibility, and Independence*, MIT Press.
- [7] Cukierman, Alex, and Allan H. Meltzer (1986), "A Theory of Ambiguity, Credibility, and Inflation under Discretion and Asymmetric Information," *Econometrica* 54, 1099-1128.
- [8] Faust, Jon (1996), "Whom Can We Trust to Run the Fed? Theoretical Support for the Founders' Views," *Journal of Monetary Economics* 37, 267-283.
- [9] Faust, Jon, and Lars E.O. Svensson (1998), "Optimal Transparency in Monetary Policy," in preparation.
- [10] Federal Reserve Bank of Kansas City (1996), *Achieving Price Stability*, Federal Reserve Bank of Kansas City Symposium Series.
- [11] Federal Reserve Board (1989), *Transcripts of Federal Open Market Committee*.
- [12] Fischer, Stanley (1990), "Rules Versus Discretion in Monetary Policy," in Benjamin M. Friedman and Frank H. Hahn, eds., *Handbook of Monetary Economics*, vol II, North Holland, Amsterdam and New York, 1155-1184.
- [13] Fuhrer, Jeffrey (1994), "Optimal monetary policy and the sacrifice ratio," in *Goals, guidelines, and constraints facing monetary policymakers*, Jeffrey Fuhrer, ed., Federal Reserve Bank of Boston, 43-69.
- [14] Goodfriend, Marvin (1986), "Monetary Mystique: Secrecy and Central Banking," *Journal of Monetary Economics* 17, 63-92
- [15] Haldane, Andrew, ed. (1995), *Targeting Inflation*, Bank of England, London.
- [16] Harvey, Andrew C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press, Cambridge.
- [17] King, Mervyn (1997), "Changes in UK Monetary Policy: Rules and Discretion in Practice," *Journal of Monetary Economics* 39, 81-98.
- [18] Kydland, Finn, and Edward Prescott (1977), "Rules rather than Discretion: the Inconsistency of Optimal Plans," *Journal of Political Economy* 85, 473-91.

- [19] Leiderman, Leonardo, and Lars E.O. Svensson, eds. (1995), *Inflation Targets*, CEPR, London.
- [20] Levine, Paul L., and Joseph G. Pearlman (1994), "Credibility, Ambiguity and Asymmetric Information with Wage Stickiness," *The Manchester School* 62, 21-39.
- [21] Lewis, Karen (1991), "Why Doesn't Society Minimize Central Bank Secrecy," *Economic Inquiry*, 29, 403-415.
- [22] McCallum, Bennett T. (1997), "Crucial Issues Concerning Central Bank Independence," *Journal of Monetary Economics* 39, 99-112.
- [23] McCallum, Bennett T. (1983), "On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective," *Journal of Monetary Economics* 11, 139-168.
- [24] Persson, Torsten, and Guido Tabellini (1990), *Macroeconomic Policy, Credibility and Politics*, Harwood, London.
- [25] Persson, Torsten, and Guido Tabellini (1993), "Designing Institutions for Monetary Stability," *Carnegie-Rochester Conference Series on Public Policy* 39, 53-84.
- [26] Rogoff, Kenneth (1989), "Reputation, Coordination and Policy," in Robert Barro, ed., *Modern Business Cycle Theory*, Harvard University Press, Cambridge, MA.
- [27] Rogoff, Kenneth (1985), "The Optimal Degree of Commitment to a Monetary Target," *Quarterly Journal of Economics* 100, 1169-1190.
- [28] Svensson, Lars E.O. (1997a), "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets," *European Economic Review* 41, 1111-1146.
- [29] Svensson, Lars E.O. (1997b), "Optimal Inflation Targets, 'Conservative' Central Banks, and Linear Inflation Contracts," *American Economic Review* 87, 98-114.
- [30] Walsh, Carl (1995), "Optimal Contracts for Central Bankers," *American Economic Review* 85, 150-167.
- [31] Walsh, Carl (1997), "Accountability, Relative Performance Measures, and Inflation Targeting," Working Paper.