

NBER WORKING PAPER SERIES

OPTIMAL MONETARY POLICY INERTIA

Michael Woodford

Working Paper 7261

<http://www.nber.org/papers/w7261>

NATIONAL BUREAU OF ECONOMIC RESEARCH

1050 Massachusetts Avenue

Cambridge, MA 02138

July 1999

I would like to thank Alan Blinder, Gregory Chow, Ray Fair, Mark Gertler, Marvin Goodfriend, Bob Hall, Pat Kehoe, Nobuhiro Kiyotaki, Phillipe Moutot, Athanasios Orphanides, Tom Sargent, Lars Svensson, John Vickers, Carl Walsh, Julian Wright, and especially Julio Rotemberg, for helpful discussions, and Marc Giannoni for excellent research assistance. I also thank the National Science Foundation, the John Simon Guggenheim Foundation, and the Center for Economic Policy Studies, Princeton University for research support. The views expressed herein are those of the authors and not necessarily those of the National Bureau of Economic Research.

© 1999 by Michael Woodford. 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.  
Optimal Monetary Policy Inertia  
Michael Woodford  
NBER Working Paper No. 7261  
July 1999  
JEL No. E52

### **ABSTRACT**

This paper considers the desirability of the observed tendency of central banks to adjust interest rates only gradually in response to changes in economic conditions. It shows, in the context of a simple model of optimizing private-sector behavior, that such inertial policy can be optimal. The reason is that small but persistent changes in short-term interest rates in response to shocks allow a larger effect of monetary policy on long rates and hence upon aggregate demand, for a given degree of overall interest-rate variability.

The paper also considers two ways of achieving the desirable degree of inertia in the equilibrium responses to shocks. One is by assignment of a loss function that penalizes squared interest-rate changes (despite the fact that interest-rate changes do not affect the true *social* objective) to a central bank that is then expected to use discretion in the pursuit of the goal. The second is through commitment to an explicit instrument rule, a generalization of the “Taylor rule” in which the funds rate is an increasing function of the lagged funds rate, as in estimated Fed reaction functions.

Michael Woodford  
Department of Economics  
Princeton University  
Princeton, NJ 08544  
and NBER  
woodford@princeton.edu

# 1 Central Bank Inertia and Optimal Policy

Many students of central bank behavior have commented on the fact that the level of nominal interest rates in the recent past appears to be an important determinant of where the central bank will set its interest-rate instrument in the present. Changes in observed conditions, such as in the rate of inflation or in the level of economic activity, result in changes in the level of the central bank's operating target for the short-term interest rate that it controls, but these changes typically occur through a series of small adjustments in the same direction, drawn out over a period of months, rather than through an immediate once-and-for-all response to the new development. This type of behavior is especially noticeable in the case of the Federal Reserve in the U.S., but characterizes many other central banks to at least some extent as well.

One way that this *inertial* character of central bank behavior shows up is in estimated central-bank reaction functions. Many studies model the Fed's behavior in terms of an implicit "target" level for the federal funds rate that depends upon the current rate of inflation and a measure of real activity.<sup>1</sup> A typical specification is of the form

$$\bar{r}_t = r^* + \phi_\pi(\pi_t - \pi^*) + \phi_y(y_t - y_t^*), \quad (1.1)$$

where  $\bar{r}_t$  is the funds rate target,  $\pi_t$  is a measure of inflation,  $\pi^*$  is the Fed's (implicit) target rate of inflation,  $y_t$  is a measure of output,  $y_t^*$  a measure of trend or potential output, and  $r^*$ ,  $\phi_\pi$ , and  $\phi_y$  are a set of constant coefficients. Estimated reaction functions of this form, however, always incorporate some form of partial adjustment of the actual funds rate toward this target, for example by specifying that the actual funds rate<sup>2</sup>  $r_t$  follows a law of motion of the form

$$r_t = \theta r_{t-1} + (1 - \theta)\bar{r}_t \quad (1.2)$$

---

<sup>1</sup>Such specifications have been used in macroeconomic models since at least the work of Fair (1978, 1979). Interest in such specifications has recently been revived by the influential discussion by Taylor (1993).

<sup>2</sup>Actually, in this equation  $r_t$  refers to the Fed's *operating target* for the funds rate, which it seeks to achieve through its daily interventions. The actual (spot) funds rate may vary considerably from this level during the day, though such deviations from the Fed's operating target are largely eliminated by the next day. See Rudebusch (1995).

where the coefficient  $\theta$  measures the degree of inertia in the central bank's response. Sack (1998b) estimates a reaction function of the form given by (1.1) – (1.2), using quarterly data for the U.S. during the Greenspan chairmanship of the Fed. His estimated value for  $\theta$  is .63, with a standard error of .08.<sup>3</sup> The estimated degree of inertia is thus large and highly significant. Other estimates for the U.S. yield similar magnitudes,<sup>4</sup> and Clarida *et al.* (1998b) report similar values of  $\theta$  in estimated reaction functions for several other central banks.

Another sort of evidence, not dependent upon an assumed specification of the variables to which the Fed reacts, is provided by Rudebusch (1995), who analyzes the statistical properties of changes in the Fed's operating target for the funds rate over the periods 1974-1979 and 1984-1992.<sup>5</sup> Rudebusch shows the Fed's discrete target changes are much more likely to be followed by another change of the same sign than by a change of the opposite sign, though the hazard rate for another change of the same sign falls as the length of time since the last target change increases, so that after an interval of five weeks with no change, the next change is about equally likely to have either sign.<sup>6</sup> Goodhart (1996) supplies further evidence in the same vein, and shows that several other central banks appear to behave similarly, though the tendency for target changes to be followed by further changes in the same direction is strongest for the Fed. These statistics are consistent with a discretized version of a partial-adjustment model like (1.2), in which a movement of the underlying target rate  $\bar{r}_t$  away from the current funds rate level results in a series of changes in the operating target  $r_t$ , each in the direction that reduces the discrepancy between  $r_t$  and  $\bar{r}_t$ .

What should one make of this feature of central bank behavior? A common view is that it shows that central banks are too slow to respond to new information that indicates

---

<sup>3</sup>This is his OLS estimate. An instrumental variables estimate is also reported, of .65 with a standard error of .06.

<sup>4</sup>See, e.g., , Clarida *et al.* (1998a) and Orphanides (1997). Fair (1978, 1979) finds an even higher degree of inertia using data from an earlier period.

<sup>5</sup>These periods are studied as they are ones in which it is reasonably clear that Fed policy was conducted in terms of an operating target for the funds rate. The target rate series is obtained from internal documents of the Federal Reserve Bank of New York Trading Desk. This series refers to the variable denoted  $r_t$  in (1.2), not to the unobserved variable  $\bar{r}_t$ .

<sup>6</sup>Sack (1998a) shows that similar conclusions hold for a sample that is extended through 1997.

the inappropriateness of current policy. The fact that further interest rate changes are somewhat forecastable is taken to indicate that the central bank delays taking actions the need for which could already be seen weeks or months earlier.<sup>7</sup> This is often supposed to be due to an unwillingness of central banks to act before the situation becomes dire enough for public opinion and political leaders to agree upon the need for action, or perhaps to a simple unwillingness of central bankers to suggest that previous policy choices can ever have been mistaken. In either case, the inertial character of policy is regarded as something that makes monetary policy less effective, especially in the stabilization of short-run fluctuations in inflation or other target variables.

It is sometimes said that central banks seek to “smooth” interest rates, in the sense that they seek to minimize the variability of interest-rate changes, in addition to their other objectives such as inflation stabilization. But models of optimizing central bank behavior that incorporate such an objective seem to be motivated mainly by a desire to rationalize the observed inertial character of interest rates, rather than by any plausible account of why such an objective is actually appropriate.<sup>8</sup> There are a number of reasons why policymakers should prefer policies that do not require the *level* of short-term interest rates to be too variable. On the one hand, the zero nominal interest-rate floor (resulting from the availability of cash as a riskless, perfectly liquid zero-return asset) means that rates *cannot* be pushed below zero. This means that a policy consistent with a low average rate of inflation, which implies a low average level of nominal interest rates, cannot involve interest-rate reductions in response to deflationary shocks that are ever too large. And at the same time, high nominal interest

---

<sup>7</sup>An explicit argument to this effect is offered by Goodhart (1998). If one supposes – as do many discussions of “inflation targeting”, especially as currently practiced in the U.K. – that the interest-rate instrument should respond solely to deviations of the central bank’s forecast of inflation some years in the future from a target rate, then in the absence of changes in the inflation target, changes in short-term interest rates should reflect changes in the inflation forecast, which should themselves be unforecastable. This view of what “inflation targeting” should entail is, however, subjected to important qualifications in Svensson (1999). The view that the desired level of interest rates at any point in time follows a random walk is also explicit, for example, in Guthrie and Wright (1998).

<sup>8</sup>We do provide a possible rationalization for a central bank “smoothing objective” below, in section 4, in terms of an optimal delegation problem. What we mean to argue here is that such a goal has no place in a true *social* objective function, so that a reason for the central bank to care about it must be sought elsewhere.

rates always imply distortions, as resources are wasted on unnecessary efforts to economize on cash balances. Friedman (1969) stresses that this is a reason to prefer a regime with low average inflation, or even moderate deflation; but it is actually the level of nominal interest rates that directly determines the size of the distortion, and the argument applies as much to short-run variation in nominal interest rates as to their average level.<sup>9</sup> Thus it is also desirable on this ground for policy not to raise nominal interests too much in response to inflationary shocks. In fact, if one supposes (by analogy with the standard argument for distorting taxes) that the distortions associated with positive nominal interest rates are a convex function of the interest rate, then, for any contemplated average level of interest rates, a lower *variance* of interest rates will reduce the average size of these distortions.<sup>10</sup> But while it makes a great deal of sense for a central bank to seek to achieve its other aims in a way consistent with as low as possible a variance of the *level* of short-run nominal rates, this in no way implies a direct concern with the variability of interest-rate *changes*.

If we discount any rational “smoothing” objective, it is often supposed that an optimizing central bank should condition its actions *only* upon state variables that affect either the current or future determination of its goal variables (such as inflation). Thus past interest rates as such should be of no significance in determining the optimal current level of interest rates, and neither should past perceptions of the direction in which interest rates needed (at that time) to be adjusted. Nor should the past behavior of goal variables (such as inflation) matter for current policy, except insofar as they may enter into central bank estimates of relevant current states or forecasts of relevant future states. Thus it is supposed that optimal policy would involve no true element of inertia at all; any persistence in interest rate fluctuations that would be observed under an optimal regime would have to be attributable

---

<sup>9</sup>See Woodford (1990) for discussion both of the generality of the argument that a positive nominal interest rate implies a distortion of private incentives, and of its implications for the desirability of low interest-rate variability.

<sup>10</sup>This is the basis for Mankiw’s (1987) argument for the desirability of interest-rate “smoothing,” in the sense of eliminating predictable interest-rate variation. But “smoothing” in Mankiw’s sense would be something very different from the sort of inertial behavior discussed here, which implies predictable interest-rate changes. In particular, Mankiw’s argument provides no reason why past interest rates, or other past states, should matter for the determination of the optimal current level of interest rates.

to serial correlation in the underlying disturbances to which the central bank responds, and not to any inertia in the central bank's own response to those disturbances.<sup>11</sup>

But this conclusion, I believe, depends upon a misunderstanding of the kind of optimal control problem that a central bank faces. In particular, it fails to take proper account of the forward-looking character of private sector behavior. In the case of a standard (engineering) optimal control problem, in which the system to be controlled evolves as a mechanical function of its current state (including possible exogenous random disturbances) and the current setting of the control variable, then it is indeed true that the controller's optimal action (from the point of view of minimization of a loss function that depends in a time-separable way upon the system state) at any point in time depends *only* upon the system's state at that time. Here the concept of the "state" of the system includes all information available at a given time that can help to forecast its future evolution (conditional upon any assumed path for the control variable), but does not include any details of past actions or states that do not continue to exert a causal influence upon the determination of goal variables at current or future dates. This is an important principle, and is at the heart of "dynamic programming" approaches to optimal control. It is also an idea that continues to pervade discussions of optimal monetary policy, which often assume a control problem of this kind, owing to the absence of forward-looking elements in many of the econometric models used for policy evaluation exercises, despite the warning of Lucas (1976) about the errors to which this can lead.

A central bank that recognizes that private sector behavior is forward-looking (as this is an inevitable result of private sector optimization) should instead realize that the evolution of its goal variables depends not only upon its current actions, but also upon how the

---

<sup>11</sup>Of course, it would be possible for an estimated central bank reaction function to include lagged interest-rate terms even in the absence of true inertia, due to serial correlation in shocks, as in the results of Aoki (1998). It is also possible that inertia might exist, consistent with the conventional ("dynamic programming") reasoning described here, simply because lagged interest rates affect the determination of current and future output and inflation, as in the results of Fair and Howrey (1996). In the model used here, such effects of lagged interest rates are excluded, to make clear that the explanation of inertia offered here is of a different kind. Note that delayed effects of monetary policy disturbances can be explained without relying upon any such effects (Rotemberg and Woodford, 1997).

private sector *expects* monetary policy to be conducted in the future. It follows from this that a more desirable outcome may be achieved if it can be arranged for private sector expectations of future policy actions to adjust in an appropriate way in response to shocks. But it does not make sense for the bank to suppose that it can manipulate expectations through announcements of intentions that bear no relation to what it actually later does. Instead, making use of this dimension upon which the economy's responses to shocks may be improved depends upon the *credibility* of the central bank's *commitments* to behave in a certain way in the future as a result of the shocks that have occurred earlier. But this credibility, which ultimately frees the central bank to act more effectively, can be maintained only if *in the short run* the bank regards itself as constrained to fulfill previous (explicit or implicit) commitments. And the need to fulfill such commitments means that central bank behavior should *not* depend solely upon current conditions and the bank's current forecast of future conditions – it should depend upon past conditions as well, and specifically those past actions by the central bank itself which (under an optimal regime) have given rise to private sector expectations of a particular kind of monetary policy now.

One reason that this critique of the traditional conception of the central bank's problem has not been more widely absorbed may have been the emphasis, in many early expositions, upon simple policy rules (such as a constant-growth-rate rule for a monetary aggregate) as the alternative to old-fashioned optimal control. Insofar as central banks have not been persuaded that judicious responses to economic developments cannot improve economic performance, their research staffs have been reluctant to abandon the study of desirable feedback rules. But, while simple rules have certain practical advantages owing to their transparency, it is *not* generally true that the optimal form of central bank commitment, taking into account the forward-looking character of private-sector behavior, is a rule that involves no adjustment of the central bank's instrument in response to current observations. Nor need the optimal commitment involve any reference to monetary targets; it may quite naturally be formulated as a feedback rule for a nominal interest-rate operating target, as the analysis below illustrates.

Another reason that the failure of the dynamic-programming principle in the case of the optimal control of forward-looking systems is not more widely appreciated is probably the popularity of analyses of optimizing central bank behavior that assume *discretionary* central bank behavior, *i.e.*, an inability of the central bank to commit itself to any future actions other than those that should appear optimal at that time. Under this assumption,<sup>12</sup> even when forward-looking elements are present in one's model, the central bank chooses its action at any given date under the assumption that it cannot affect private sector expectations about its behavior at later dates, owing to its inability to commit itself. Hence its optimization problem at each date can be cast in a dynamic-programming form, and the resulting solution makes its action a function only of the economy's current state.

But even if this model of central bank behavior involves *optimization*, it does not represent *optimal policy*. That is to say, it does not achieve the best possible outcome, in terms of the central bank's objective function, among those that could be achieved by an appropriate rule for central bank behavior. The celebrated analysis by Barro and Gordon (1983) of the undesirably high average rate of inflation that results under discretion illustrates the suboptimal nature, in general, of what is effectively a noncooperative equilibrium of a game played by a succession of policymakers. But while the Barro-Gordon point is widely appreciated, many discussions seem to presume that the problem with discretionary policymaking is solved once one substitutes an alternative target for the average rate of inflation (a central banker that desires lower inflation than is truly optimal) or for the average output gap (a central banker that does not seek to keep output above the natural rate, even though the natural rate is inefficiently low), or perhaps an alternative weighting of the bank's stabilization objectives (a central banker that cares more about inflation stabilization than does society). Instead, there is typically a similar problem with discretionary policymaking in regard to the optimal response of policy to random *shocks*, even if one has adjusted the central bank's objective so as to make discretion compatible with an optimal steady state in the absence

---

<sup>12</sup>Here I have in mind the common practice of considering only the Markov-perfect solution of such a model. See further discussion in sections 3.2 and 4.1 below.

of shocks. Here, too, an optimal *commitment* on the part of the central bank can often lead to a better outcome than is attained through optimization under discretion.<sup>13</sup>

The essential insight into why a commitment to inertial behavior may be optimal for a central bank is provided by a suggestion of Goodfriend (1991) regarding the reason for apparent interest-rate “smoothing” by banks like the Fed.<sup>14</sup> Goodfriend argues that output and prices do not respond to daily fluctuations in the (overnight) federal funds rate, but only variations in *longer-term* interest rates. The Fed can thus achieve its stabilization goals only insofar as its actions affect these longer-term rates. But long rates should be determined by market expectations of future short rates. Hence an effective response by the Fed to inflationary pressures, say, requires it to communicate a credible commitment to a changed *future path* of short rates. One straightforward way to do this is to establish a reputation for maintaining interest rates at a higher level for a period of time once they are raised – or even for following initial small interest-rate changes by further changes in the same direction, in the absence of a change in conditions that makes this unnecessary. Such a policy, if understood by the private sector, offers the prospect of significant effects of central bank policy upon aggregate demand, without requiring excessively volatile short-term interest rates, which would be undesirable for the reasons summarized earlier.

There is a certain amount of evidence suggesting that in the U.S. at least, the inertial character of Fed policy has this beneficial effect. Cook and Hahn (1989), Rudebusch (1995), and Goodhart (1996) all present evidence showing that changes in Fed operating targets for the federal funds rate do affect longer-term interest rates, especially at the time of changes in the direction of movement of the target. Furthermore, Watson (1999) argues that the increased volatility of long rates since the mid-1980s (relative to an earlier 1960-75 sample period) can be explained by a greater degree of persistence of funds rate movements in the later period – a period that has been characterized by greater success at inflation stabilization, and, many would argue, an increase in the credibility of the Fed’s commitment to

---

<sup>13</sup>Clarida *et al.* (1999) also discuss this point in the context of a similar model.

<sup>14</sup>This explanation is also endorsed by Rudebusch (1995).

a consistent anti-inflationary policy. Indeed, some commentators have proposed that U.S. monetary policy has been so successful at inflation stabilization in the 1990s, despite relatively little change in the funds rate for years at a time, because “the bond market does the Fed’s work for it,” responding to disturbances in the way needed to keep inflation stable without the need for large policy adjustments by the Fed. This is exactly what a good policy regime *should* look like, according to the analysis that I shall offer here – not because the bond market has any *reason* to react, of course, if short rates will never be adjusted at all, but because a credible commitment to an optimal (highly inertial) feedback rule on the part of the Fed should not require large movements of short-term interest rates in equilibrium, highly persistent low-amplitude variations being sufficient to achieve a desirable degree of inflation stabilization.

This interpretation of the nature of monetary policy inertia has the advantage of resolving the mystery of why central banks that exhibit a great deal of inertia in their behavior (such as the Fed) nonetheless have quite good records with respect to inflation control – an observation that is mostly puzzling under the view that inertia is an obstacle to effective central bank action. Whether one should suppose that these banks have been best able to control inflation because of their better understanding of the advantages of inertial behavior, or whether it is rather that the advantages of inertial behavior only become important in the case of central banks that have first established a certain degree of credibility as a result of the consistency of their policy, I shall not here seek to judge. But it makes a great deal of sense that those central banks that have most clearly learned the benefits of commitment with regard to the average rate of inflation should also be the ones that are also most able to benefit from a perceived commitment to predictable responses to shocks as well – a commitment that should manifest itself in inertial behavior. <sup>15</sup>

---

<sup>15</sup>An alternative explanation, that would also make gradual policy changes optimal, without requiring a rejection of the dynamic-programming principle, attributes policy inertia to gradual central bank learning about the effects of its actions (Caplin and Leahy, 1996; Sack, 1998a, 1998b). The reason offered here for optimal policy to be inertial is quite different. In particular, I shall assume a linear-quadratic approximation to the central bank’s optimization problem, under which optimal policy is *certainty-equivalent*. This means that a predictable increase over time in the precision of the central bank’s estimate of the current state should not be a reason for forecastable changes in its desired instrument setting.

The analysis below offers a formal analysis of the benefits of inertial behavior in the context of a simple, and now rather standard, forward-looking macro model, with clear foundations in optimizing private sector behavior. The model is too simple to provide the basis for a realistic quantitative policy analysis; a companion paper (Rotemberg and Woodford, 1998) that addresses this issue in the context of a small econometric model fit to U.S. time series is perhaps more interesting in that regard. But numerical results in the context of a specific quantitative model, of the sort offered in the other paper, inevitably raise many questions about how sensitive the conclusions may be to a long list of arguable assumptions that are made along the way. The analytical treatment here of a simplified version of that model is intended to help clarify the source of some of those numerical results, and thus to suggest lessons that may be of more general validity.

The paper is organized as follows. Section 2 presents the model of the economy and poses the problem of optimal monetary policy. Section 3 characterizes the responses of endogenous variables, including nominal interest rates, to shocks under an optimal regime, and highlights the advantages of commitment, by contrasting the optimal responses with those that would result from optimization under discretion. It is shown that the optimal responses involve intrinsic inertia in interest-rate responses, in addition to the persistence resulting from persistence of the exogenous real disturbances themselves. Section 4 then considers the optimal assignment of an objective to a central bank with instrument (but not goal) independence, that is expected to pursue its assigned goal under discretion. It is shown that in this case, it is desirable for the central bank's loss function to include an "interest-rate smoothing" objective, even though the true social loss function does not. Finally, section 5 considers the form of interest-rate feedback rule that can achieve the desired dynamic responses to shocks, if the central bank's commitment to such a rule is credible to the private sector. It is shown that such a rule must involve dependence of the current operating target upon the level of interest rates in the recent past, as is characteristic of estimated central bank reaction functions.

## 2 The Problem of Optimal Monetary Policy

In order to illustrate more concretely the themes of the preceding discussion, it is useful to introduce a simple optimizing model of inflation and output determination under alternative monetary policies, where monetary policy is specified in terms of a feedback rule for a short-term nominal interest rate instrument. The model is similar, if not identical, to the small forward-looking models used in a number of recent analyses of monetary policy rules, including Kerr and King (1996), Woodford (1996), Bernanke and Woodford (1997), McCallum and Nelson (1997, 1998), Kiley (1998), and Clarida *et al.* (1999). As is explained in Woodford (1996), the model's equations can be derived as log-linear approximations to the equilibrium conditions of a simple intertemporal general equilibrium model with sticky prices. While the model is quite simple, it incorporates forward-looking private sector behavior in three respects, each of which is surely of considerable importance in reality, and would therefore also be present in some roughly similar form in any realistic model. It also shares many features with the econometric model of Rotemberg and Woodford (1997, 1998), and so analysis of this model can provide insight into the source of some of the numerical results obtained there.

The model's two key equations are an *intertemporal IS equation* of the form

$$y_t - g_t = E_t[y_{t+1} - g_{t+1}] - \sigma^{-1}[r_t - E_t\pi_{t+1}], \quad (2.1)$$

and an *aggregate supply equation* of the form

$$\pi_t = \kappa[y_t - y_t^n] + \beta E_t\pi_{t+1}, \quad (2.2)$$

where  $y_t$  is the deviation of the log of real output from its trend path,  $\pi_t$  is the rate of inflation (first difference of the log of the price level), and  $r_t$  is the deviation of the short-term nominal interest rate (the central bank's policy instrument) from its steady-state value in the case of zero inflation and steady output growth.<sup>16</sup> These two equations, together

---

<sup>16</sup>Thus all three variables denote percentage deviations from the values of the variables in a steady state with zero inflation and output growth at the constant trend rate.

with an interest-rate rule such as (1.1) – (1.2) representing monetary policy, determine the equilibrium evolution of the three endogenous variables  $\pi_t$ ,  $y_t$ , and  $r_t$ . The two exogenous disturbances to these structural equations are the processes  $g_t$ , representing autonomous variation in spending not motivated by intertemporal substitution in response to real interest-rate changes, and  $y_t^n$ , representing time variation in the “natural rate” of output, which would be the equilibrium level of output under perfectly flexible prices (and is independent of monetary policy). Finally, on theoretical grounds the structural parameters  $\sigma$  and  $\kappa$  are both positive.

The system of equations (2.1) – (2.2) implies that no *lagged* values of any endogenous variables play any role in the determination of the equilibrium values of inflation, output, or interest rates at a given point in time, unless such dependence is introduced by the monetary policy rule. On the other hand, they do involve important dynamic linkages from expectations of the *future* to the present, as both  $E_t\pi_{t+1}$  and  $E_ty_{t+1}$  enter the equations that determine equilibrium at date  $t$ . The complete exclusion of inertial terms in these structural equations means that if the “dynamic programming principle” referred to above were valid, the *only* source of inertia in the endogenous variables under an optimal policy would be persistence in the exogenous disturbances ( $g_t$  and  $y_t^n$ ). Analysis of such a system thus makes it especially clear how the presence of forward-looking elements in the structural equations renders that principle invalid. The model is obviously an extremely simple one, and in particular, the complete absence of inertial terms in the structural equations is not entirely realistic. But what is more important for present purposes is the justification of the forward-looking elements, that are critical for the results obtained below.

There is in fact considerable support, both theoretical and empirical, for each of the crucial forward-looking elements in this system. First of all, aggregate demand (intertemporal substitution) depends upon real as opposed to nominal interest rates, as a result of which inflation expectations enter the IS equation. Second, current aggregate demand depends not simply upon current expected *short-term* real rates of return, but also upon expected *future* aggregate demand, since (2.1) indicates the incentives for intertemporal substitution

of demand created by real interest rate expectations (Kerr and King, 1994). Alternatively, (2.1) can be integrated forward to yield

$$y_t = g_t - \sigma^{-1} \sum_{j=0}^{\infty} E_t[r_{t+j} - \pi_{t+j+1}]. \quad (2.3)$$

Under this way of writing the equation, it states that aggregate demand depends not upon current short rates alone, but rather upon expected *long-term* real rates, which in turn depend upon expected *future* short rates, as specified for example by Fuhrer and Moore (1995). This aspect of the assumed effects of monetary policy upon aggregate demand is a critical element in the argument of Goodfriend (1991), which we seek to formalize. Here it might be objected that the expectations theory of the term structure of interest rates has been subject to considerable empirical criticism. But these criticisms relate mainly to the hypothesis that long rates are moved *only* by changes in rational expectations of future short rates; they do not show that forecastable movements in future short rates do not affect long rates as the theory would predict. The inclusion of a time-varying “term premium” is not a problem for the justification of an IS relation of the form (2.3), as long as the term premium is exogenous. For then variations in the term premium are simply another source of variations in the disturbance term  $g_t$ .

Finally, because prices are not expected to be continuously changed, they are set on the basis of expectations of future cost and demand conditions, and not just on the basis of current conditions. This results in the expected inflation term in (2.2). If one adopts the device introduced by Calvo (1983) of exogenous stochastic intervals between price changes, together with monopolistic competition among price-setting suppliers, a log-linear approximation to the first-order condition for optimal price-setting implies that average prices should adjust at a rate  $\pi_t$  that satisfies

$$\pi_t = \omega c_t + \beta E_t \pi_{t+1}, \quad (2.4)$$

where  $c_t$  represents (percentage deviations in) the average level of real marginal cost of supply, the coefficient  $0 < \beta < 1$  represents the factor with which suppliers discount future real income<sup>17</sup>, and the coefficient  $\omega > 0$  depends upon the average frequency of price changes and

the elasticity of demand faced by suppliers of individual goods. Prices are more nearly flexible the higher is  $\omega$ , as a high  $\omega$  indicates a low tolerance for discrepancies between one's current price (or more precisely, marginal revenue, which is proportional to one's price) and one's current nominal marginal cost. If one writes (2.4) as an equation to determine the current average price level as a weighted average of the past average price level, current average marginal cost, and the expected future average price level, then as  $\omega$  is made unboundedly large, the weights on both past and future prices go to zero. If one adjoins to this a simple theory of cyclical variation in real marginal cost of the form

$$c_t = \eta[y_t - y_t^n], \tag{2.5}$$

then one obtains an aggregate supply relation of the form (2.2). Here the exogenous disturbance term  $y_t^n$  represents the level of output consistent with  $c_t = 0$  at each point in time, and hence the equilibrium level of output in the limiting case of perfect price flexibility.

Empirical support for this sort of forward-looking model of pricing behavior is found, for example, in Roberts (1995), Sbordone (1998), and Galí and Gertler (1998). Sbordone tests the implications of (2.4) using average unit labor costs as a proxy for the average level of nominal marginal cost, and shows that the path of inflation implied by the evolution of unit labor costs given (2.4), and using a VAR model to forecast future growth in unit labor costs, is quite similar to the actual path of U.S. inflation. Furthermore, she shows not only that pricing equation (2.4) fits the data much better than would the hypothesis of flexible prices, but that the presence of the forward-looking term considerably improves the fit of the model. Specifically, nonlinear least-squares estimation of the parameters  $\omega$  and  $\beta$  implies that  $\beta$  is significantly greater than zero,<sup>18</sup> and in fact quite close to one, in addition to implying a significantly positive value for  $1/\omega$ . Galí and Gertler similarly estimate the pricing equation (2.4), using a slightly different measure of marginal cost and an instrumental-variables technique, and reach similar conclusions. When they generalize

---

<sup>17</sup>This corresponds to the discount factor of the representative household, in a general-equilibrium version of the model, as in Woodford (1996).

<sup>18</sup>The case of  $\beta = 0$ , but  $\omega$  finite, would correspond to a model in which prices are set proportional to a weighted average of current and past unit labor costs, with no forward-looking terms.

the model to allow for a mixture of forward-looking and backward-looking price-setters, they estimate a larger fraction of the population to obey the forward-looking pricing equation, and find that the presence of the backward-looking price-setters accounts for relatively little of aggregate inflation dynamics.

Roberts (1995) instead directly estimates the “new-Keynesian Phillips curve” (2.2), using a variety of proxies for inflation expectations and for variations in the natural rate of output. Most of his results similarly support the existence of a significantly positive coefficient on  $E_t\pi_{t+1}$ , and the poorer fit obtained in this case may relate more to the simplicity of the model of marginal cost (2.5), or to failure to adequately control for variation in  $y_t^n$ , than to any problem with the theory of pricing. Roberts (1998) suggests instead that the aggregate supply relation fits better if the assumption of rational expectations is replaced by a partially backward-looking model of expected inflation. But even here, his preferred specifications (partial adjustment of expectations *toward* the rational forecast) imply important forward-looking elements in the AS relation of the kind specified in (2.2).

It is convenient to rewrite the model’s structural equations in terms of the *output gap*  $x_t \equiv y_t - y_t^n$ . Equation (2.2) may then be simply written

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1}. \quad (2.6)$$

Since we shall assume that we are interested in the output gap rather than detrended output anyway, in evaluating the success of stabilization policy, it is useful to rewrite the equation in this way, that eliminates the disturbance term. Equation (2.1) correspondingly becomes

$$x_t = E_t x_{t+1} - \sigma^{-1} [(r_t - r_t^n) - E_t \pi_{t+1}] \quad (2.7)$$

where

$$r_t^n \equiv \sigma E_t [(y_{t+1}^n - y_t^n) - (g_{t+1} - g_t)].$$

The exogenous disturbance  $r_t^n$  corresponds to Wicksell’s “natural rate of interest”, the interest rate (determined by purely *real* factors) that would represent the equilibrium real rate of return under flexible prices, and that corresponds to the nominal interest rate consistent

with an equilibrium with constant prices.<sup>19</sup> In our simple model, disturbances to the natural rate represent a useful summary statistic for *all* non-monetary disturbances that matter for the determination of inflation and the output gap, for no other disturbance term enters either equation (2.6) or (2.7). Hence if, as we shall suppose, the goals of stabilization policy can be described in terms of the paths of the inflation rate, the output gap, and interest rates alone, then the problem of optimal monetary policy may be formulated as a problem of the optimal response to disturbances to the natural rate of interest. It may furthermore be observed that neither the interest rate controlled by the central bank nor the natural rate enters either of the structural equations, *except* through the “interest-rate gap”  $r_t - r_t^n$ . Thus non-monetary disturbances matter only insofar as the interest rate controlled by the central bank fails to track the resulting fluctuations in the natural rate; and whether the interest rate set by the central bank at any point in time should be considered high or low, in the sense that is relevant for inflation determination, depends purely upon where it is *relative* to the current natural rate.

I shall assume that the objective of monetary policy is to minimize the expected value of a loss criterion of the form

$$W = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t L_t \right\}, \quad (2.8)$$

where  $0 < \beta < 1$  is a discount factor, and the loss each period is given by

$$L_t = \pi_t^2 + \lambda_x (x_t - x^*)^2 + \lambda_r (r_t - r^*)^2, \quad (2.9)$$

for some optimal levels  $x^* \geq 0$  and  $r^*$  of the output gap and the nominal interest rate, and some weights  $\lambda_x, \lambda_r > 0$ . The assumed form of (2.9) is relatively conventional, except that an interest-rate variability term is included, for either or both of the reasons discussed in the introduction. (Note that an interest-rate “smoothing” objective is *not* assumed!) The sign of  $r^*$  depends upon whether it is desirable to stabilize nominal interest rates around a value greater or less than the level associated with a steady-state equilibrium with zero

---

<sup>19</sup>See Blinder (1998, chap. 2) for a recent discussion of the usefulness of this concept in the theory of central banking.

inflation. Insofar as the interest-rate term represents the distortions emphasized by Friedman (1969), one would expect  $r^*$  to be a small *negative* quantity (since a zero nominal interest rate corresponds to  $r^* = \log \beta < 0$  in our notation). But insofar as it represents an approximation to the constraint imposed by the zero bound on nominal interest rates, it may be appropriate to set  $r^* > 0$ , as a higher average value of interest rates reduces the extent to which interest-rate variability is constrained by the zero floor.<sup>20</sup>

The allowance for  $x^* > 0$  reflects the fact that the natural rate of output may well be inefficiently low. The explicit general equilibrium model underlying the structural equations presented here (as presented, for example, in Woodford, 1996) implies that this should be so, as a result of the small amount of market power that the producers of differentiated goods are assumed to have (in order to allow them the power to set prices), if one does not assume an offsetting output subsidy.<sup>21</sup> Finally, I assume that the distortions associated with inflation are minimized in the case of zero inflation. Some might prefer to instead see a term of the form  $(\pi_t - \pi^*)^2$  in (2.9), where  $\pi^*$  is not necessarily zero. This would in fact make little difference for our conclusions below.<sup>22</sup> But there would seem to be good theoretical grounds to argue that the distortions associated with price-level instability are minimized when prices never change (rather than by some other steady inflation rate). And some commonly discussed reasons to prefer a non-zero average inflation rate are appropriately modeled by a non-zero value for  $r^*$ , rather than a non-zero  $\pi^*$ ; these include both Friedman's (1969) argument for mild deflation and Summers' (1991) argument for mild inflation, since both arguments really are about the desirable average level of nominal interest rates.

---

<sup>20</sup>The derivation of the interest-rate variation term in the loss function from either of these two concerns is taken up in Woodford (1999b, sec. 4), which also explains the theoretical values of both  $\lambda_r$  and  $r^*$ .

<sup>21</sup>The welfare calculations in Rotemberg and Woodford (1997, 1998) are carried out under the assumption that  $x^* = 0$ , in order to focus on the welfare losses associated with price stickiness and imperfect stabilization of the effects of shocks, in abstraction from other distortions that it is not properly the job of monetary policy to address. However, whether this is assumed or not has little effect upon our conclusions about the nature of optimal responses to shocks, as will be seen. It matters much more for the analysis of discretionary monetary policy, as shown in the next section.

<sup>22</sup>It would affect the conclusion reached about the steady-state rate of inflation that should be implicit in the optimal policy rule, for obvious reasons, but it would have no effect upon our conclusions about the desirable responses to shocks, or the desirable type of feedback from endogenous variables to interest rates.

Finally, it should be noted that the presence of an  $(x_t - x^*)^2$  term in (2.9) assumes that the random fluctuations in the natural rate of output (shifts in the aggregate supply curve) are all variations in the *efficient* level of output as well, even if the efficient level of output may at all times be a certain number of percentage points above the natural rate. This makes sense if the distortions resulting from delays in price adjustment, on the one hand, and from a constant level of market power, on the other, are the only reasons why equilibrium output ever differs from the efficient level. In this case, disturbances to the natural rate of output due to variations in government purchases, exogenous changes in household tastes, or random variation in the rate of technical progress, will indeed shift the natural rate and the efficient level of output in the same proportion (in our log-linear approximation to the rational expectations equilibrium). But other sorts of distortions might, in principle, result in time variation of different sorts in the two series.<sup>23</sup> Such complications would require an additional term in (2.9).

An important justification for the specific form of objective assumed here is that it represents a second-order Taylor series approximation to the theoretically correct welfare measure, the expected utility level of the representative household, in the same optimizing model as is used to derive the structural equations (2.1) – (2.2).<sup>24</sup> Let us suppose that the fluctuations in the natural rate of interest  $r_t^n$  around its average value satisfy a uniform bound that is linear in a parameter  $\|\xi\|$ , and let us suppose that the coefficients  $x^*$  and  $r^*$  of the loss function (2.9), that measure the degree of inefficiency of the zero-inflation steady state, also satisfy bounds that are linear in  $\|\xi\|$ . Then our log-linear approximate structural equations (2.6) – (2.7) are accurate approximations to the exact, nonlinear equilibrium conditions implied by the optimizing model in the case that  $\|\xi\|$  is made small enough; strictly speaking, we can show that the distance between the paths of the endogenous variables in the solution to these approximate structural equations and in the exact equilibrium is of order  $\mathcal{O}(\|\xi\|^2)$

---

<sup>23</sup>This is the point, for example, of the assumption by Clarida *et al.* (1999) of a “cost-push” disturbance term in (2.2) that is not netted out in their definition of the “output gap” that policymakers are assumed to wish to stabilize.

<sup>24</sup>See Woodford (1999b) for details of the derivation.

or smaller.<sup>25</sup> The welfare measure (2.8) with period loss function

$$\tilde{L}_t = \pi_t^2 + \lambda_x(x_t - x^*)^2 \quad (2.10)$$

represents a quadratic approximation to the exact theoretical welfare measure (expected utility) in the same sense; the difference between  $W$  and the exact level of expected utility<sup>26</sup> is of order  $\mathcal{O}(\|\xi\|^3)$ .

Such a quadratic approximation is appropriate, not only to facilitate comparison between our results and those elsewhere in the literature, but because this is in any event the highest-order approximation to utility that can be computed using only our log-linear approximate characterization of the equilibrium resulting from a given policy rule.<sup>27</sup> Furthermore, a quadratic approximation to the theoretically correct objective suffices to allow us to obtain a log-linear approximation to the optimal responses to shocks, by finding the policy that minimizes our approximate loss criterion subject to the log-linear constraints imposed by our approximate structural equations.<sup>28</sup> Since studies that characterize optimal policy on the basis of the exact utility functions and first-order conditions (such as King and Wolman, 1998) frequently use log-linear approximations to characterize the optimal responses to shocks in any event, this order of accuracy is as much as we really need.

The final term in (2.9) can be justified either as representing further distortions associated with high nominal interest rates (for example, inefficient substitution between “cash” and “credit” goods, as in Yun, 1996), or as a quadratic approximation to the penalty on interest-

---

<sup>25</sup>This result assumes the existence of a determinate equilibrium associated with the policy rule under consideration, in the sense discussed in section 4.1 below, and that that unique bounded solution is the one under discussion.

<sup>26</sup>Here we have omitted from  $W$  constant terms that are independent of the evolution of the endogenous variables, and have chosen units for the measurement of  $W$  in which the period loss associated with steady inflation of one percent per year is equal to a flow loss of one per time period.

<sup>27</sup>Even a second-order approximation to utility can be computed only because of our assumption that the parameters  $x^*$  and  $r^*$  are of order  $\mathcal{O}(\|\xi\|)$ . This means that only contributions of order  $\mathcal{O}(\|\xi\|)$  to the average levels of the endogenous variables matter for the terms of second order or larger in  $W$ . Otherwise, second-order contributions to the average level of variables such as  $x_t$  would matter, and these could not be computed without a higher-order approximation to our model structural equations.

<sup>28</sup>The first-order conditions obtained in section 3.1 below differ from the exact ones by terms that are only of order  $\mathcal{O}(\|\xi\|^2)$ . This depends upon the fact that the steady-state values of the Lagrange multipliers associated with the constraints imposed by the structural equations are only of order  $\mathcal{O}(\|\xi\|)$ .

rate variability implied by the zero bound on nominal rates. We shall emphasize the latter interpretation here. Rotemberg and Woodford (1997, 1998) propose to approximate the effects of the zero bound by imposing instead a requirement that the mean federal funds rate be at least  $k = 2.26$  standard deviations above zero.<sup>29</sup> The alternative constraint, while inexact, has the advantage that checking it requires only computation of first and second moments under alternative policy regimes; and, given linear structural equations and a quadratic loss function, the optimal policy is linear. But a constraint of this form can equivalently be expressed as a requirement that the average value of  $r_t^2$  be not more than  $K \equiv 1 + k^{-2} = 1.44$  times the square of the average value of  $r_t$ . If we use discounted averages, for conformity with the other terms in our welfare measure, we obtain a constraint of the form

$$E_0 \left[ (1 - \beta) \sum_{t=0}^{\infty} \beta^t r_t^2 \right] \leq K E_0 \left[ (1 - \beta) \sum_{t=0}^{\infty} \beta^t r_t \right]^2. \quad (2.11)$$

And using the usual Kuhn-Tucker arguments, the policy that minimizes the expected discounted value of (2.10) subject to (2.11) can be shown to also minimize an (unconstrained) loss criterion of the form

$$E_0 \left[ (1 - \beta) \sum_{t=0}^{\infty} \beta^t \tilde{L}_t \right] - \mu_1 E_0 \left[ (1 - \beta) \sum_{t=0}^{\infty} \beta^t r_t^2 \right] + \mu_2 E_0 \left[ (1 - \beta) \sum_{t=0}^{\infty} \beta^t r_t \right],$$

where  $\mu_1$  and  $\mu_2$  are appropriately chosen Lagrange multipliers. Finally, the terms in this expression can be rearranged to yield a discounted loss criterion of the form (2.8) – (2.9), where  $\lambda_r \equiv \mu_1$  and  $r^* \equiv \mu_2/2\mu_1$ .<sup>30</sup> When constraint (2.11) binds, the Lagrange multipliers have values  $\mu_1, \mu_2 > 0$ , so that  $\lambda_r, r^* > 0$ . We adopt this interpretation below, when “calibrating” the parameters of the loss function.

An advantage of deriving the loss function from welfare-economic foundations in this way is that it not only provides justification for the general form of our welfare criterion, but for a specific quantitative specification. The theoretical derivation implies, for example, that

---

<sup>29</sup>This numerical value is the one satisfied by the stationary distribution for the federal funds rate implied by their estimated VAR using U.S. data.

<sup>30</sup>There is also a constant term involved in completing the square, that has no effect upon our ranking of alternative policies.

the discount factor in (2.8) should be the same as the coefficient on expected inflation in the structural equation (2.2), as both are equal to the discount factor of the representative household in the underlying optimizing model. It similarly implies that  $\pi^*$  should equal zero in (2.9), as I have assumed in writing the equation, and implies theoretical values for coefficients such as  $\lambda_x$  and  $x^*$  in the terms of the same underlying model parameters (mainly preference parameters of the representative household) as determine the values of the coefficients  $\sigma$  and  $\kappa$  of the structural equations. These theoretical relations are used in the illustrative numerical calculations reported below. However, the algebraic characterizations of optimal policy apply for arbitrary values of the parameters of the period loss function (2.9), so our qualitative conclusions do not depend upon this particular view of the proper objectives of monetary policy, but only upon a loss function of the (rather conventional) general form (2.8) – (2.9).

### 3 Optimal Responses to Fluctuations in the Natural Rate of Interest

I turn now to the characterization of optimal monetary policy in the context of the model just set out. In the present section, I shall be concerned solely with the question of what pattern of fluctuations in the endogenous variables, inflation, output and interest rates, would be associated with the optimal equilibrium, by which I mean the equilibrium that achieves the lowest possible value of the loss measure (2.8). I shall set aside until the next section the question of what kind of policy rule should bring about such an equilibrium, were it to be properly understood by the private sector. Since the only random disturbances that matter for this optimization problem are the variations in the natural rate of interest  $r_t^n$ , the problem of this section can be alternatively posed as the question of how it is optimal for inflation, the output gap, and nominal interest rates to respond to exogenous fluctuations in the natural rate.

Formally, our problem is to choose stochastic processes  $\pi_t$ ,  $x_t$ , and  $r_t$  — specifying each of these variables as a function of a random state  $I_t$  that includes not only the complete

history of the exogenous disturbances  $(r_t^n, r_{t-1}^n, \dots, r_0^n)$ , but also all public information at date  $t$  about the future evolution of the natural rate<sup>31</sup>— in order to minimize the criterion defined by (2.8) and (2.9), subject to the constraint that the processes satisfy equilibrium conditions (2.6) and (2.7) at all dates  $t \geq 0$ . We imagine that a policymaker can choose the entire future (state-contingent) evolutions of these variables, once and for all, at date zero. Thus we wish to consider optimal policy under *commitment* on the part of the policymaker – even though we have not yet specified the type of explicit commitment, as to the way in which policy will be conducted, that is involved. Note that the assumed possibility of commitment matters, in the case of an optimization problem of this kind. For, because of the forward-looking terms in our structural equations (2.6) and (2.7), the value of the period loss  $L_t$  that can be achieved at a given time depends upon what the private sector expects about the subsequent evolution of the endogenous variables. *Commitments* about how policy will be used to affect those variables at later dates, that need not coincide with what would optimally be chosen at the later dates in the absence of an advance commitment, will thus be a typical feature of an optimal plan.

### 3.1 First-Order Conditions for the Optimal Plan

This sort of linear-quadratic optimization problem can be treated using methods that are by now familiar.<sup>32</sup> It is useful to write a Lagrangian of the form<sup>33</sup>

$$E_0 \left\{ \sum_{t=0}^{\infty} \beta^t \{ L_t + 2\phi_{1t} [x_t - x_{t+1} + \sigma^{-1}(r_t - r_t^n - \pi_{t+1})] + 2\phi_{2t} [\pi_t - \kappa x_t - \beta \pi_{t+1}] \} \right\}. \quad (3.1)$$

---

<sup>31</sup>We may at this point also allow  $I_t$  to include other information, including irrelevant “sunspot” variables, but we shall find that it is not optimal for the endogenous variables to respond to any such additional information.

<sup>32</sup>See, e.g., Backus and Driffill (1986) for treatment of a general linear-quadratic problem. Early applications of the Lagrangean method used here include Kydland and Prescott (1980), Hansen, Epple and Roberds (1985), and Sargent (1987, chap. XV). Methods of this kind are applied extensively to problems of optimal monetary policy by Currie and Levine (1993).

<sup>33</sup>Here the Lagrange multipliers are multiplied by two to eliminate a recurrent factor of two from the resulting first-order conditions; the same result is often instead achieved by defining the loss function  $L_t$  to equal one-half of (2.9). Note also that conditional expectations are dropped from the way in which the constraints are written inside the square brackets, because the expectation  $E_0$  at the front of the entire expression makes them redundant.

An optimal plan then must satisfy the first-order conditions

$$\pi_t - (\beta\sigma)^{-1}\phi_{1t-1} + \phi_{2t} - \phi_{2t-1} = 0, \quad (3.2)$$

$$\lambda_x(x_t - x^*) + \phi_{1t} - \beta^{-1}\phi_{1t-1} - \kappa\phi_{2t} = 0, \quad (3.3)$$

$$\lambda_r(r_t - r^*) + \sigma^{-1}\phi_{1t} = 0, \quad (3.4)$$

obtained by differentiating the Lagrangian with respect to  $\pi_t$ ,  $x_t$ , and  $r_t$  respectively. Each of conditions (3.2) – (3.4) must hold at each date  $t \geq 1$ , and the same conditions also must hold at date  $t = 0$ , where however one adds the stipulation that

$$\phi_{1-1} = \phi_{2-1} = 0. \quad (3.5)$$

We may omit consideration of the transversality conditions, as we shall consider only bounded solutions to these equations, which necessarily satisfy the transversality conditions.<sup>34</sup> A (bounded) optimal plan is then a set of bounded processes  $\{\pi_t, x_t, r_t, \phi_{1t}, \phi_{2t}\}$  for dates  $t \geq 0$ , that satisfy (2.6), (2.7), and (3.2) – (3.4) at all of these dates, consistent with the initial conditions (3.5).

It is obvious that such an optimal plan will, in general, not be *time consistent*, in the sense discussed by Kydland and Prescott (1977). For a policymaker that solves a corresponding problem at some date  $T > 0$  will choose processes for dates  $t \geq T$  that satisfy the stochastic difference equations just listed, *and* initial conditions

$$\phi_{1T-1} = \phi_{2T-1} = 0$$

corresponding to (3.5). But these last conditions will, in general, not be satisfied by the optimal plan (with commitment) chosen at date zero. This can be shown as follows. The optimal plan is time consistent only if  $\phi_{it} = 0$  for  $i = 1, 2$  and for all  $t \geq 0$ . Substituting these values into (3.2) – (3.4), we see that this requires that the optimal plan involve  $\pi_t = 0, x_t =$

---

<sup>34</sup>Here it is important to recall that we are not solving an *exact* linear-quadratic problem, which might have an unbounded solution. Our Taylor series methods imply that the solution to the linear-quadratic problem taken up here should approximate the solution to the exact nonlinear welfare maximization problem *only* in the case that solution involves fluctuations in the state variables that are uniformly small.

$x^*, r_t = r^*$  at all times. But substitution of these values into (2.6) and (2.7) shows that this cannot represent an equilibrium, unless  $x^* = 0$ , and  $r_t^n = r^*$  at all times. Thus the optimal plan is time consistent *only* if the problem satisfies both of those properties. The failure of time consistency in the case that  $x^* \neq 0$  (i.e., the natural rate of output is inefficient, and so not the level that a benevolent policymaker would prefer to achieve) occurs essentially for the same reason as in the celebrated analysis of Barro and Gordon (1983). However, even if we assume that  $x^* = 0$ ,<sup>35</sup> the optimal plan is still generally not time consistent, for time consistency would also require that there be no variation in the natural rate of interest.<sup>36</sup> Even in the absence of any desire to make the level of output higher than the natural rate on average, and hence of a systematic inflation bias, a more desirable pattern of responses to *shocks* is available under commitment. This is why *credibility* is valuable to a central bank, and the commitment to a *policy rule* desirable, even in the absence of such bias.

### 3.2 Comparison with the Result of Optimization without Commitment

It may be useful to compare the optimal (commitment) plan to the *time-consistent* plan that results from optimization under discretion, i.e., optimization by a central bank that expects itself to re-optimize at each successive date, and is unable to commit itself to do anything at future dates other than choose a policy that is expected to minimize (2.8) given the expected behavior of its successors. By this latter plan I shall mean the Markov-perfect equilibrium of the game played by the succession of central bankers, or what is sometimes called the “non-reputational” solution to the central bank’s problem (e.g., Currie and Levine, 1993). Under the time-consistent plan,  $\pi_t, x_t$  and  $r_t$  are chosen at any given date  $t$ , subject

---

<sup>35</sup>This might be either because the discretionary policymaker seeks to minimize a version of (2.9) in which  $x^* = 0$  despite the inefficiency of the natural rate (as assumed in Clarida *et al.*, 1999), or because the natural rate is actually assumed to be efficient (as in the welfare calculations of Rotemberg and Woodford, 1998).

<sup>36</sup>Actually, this condition is required only because we have assumed that  $\lambda_r > 0$ . If  $x^* = 0$  and  $\lambda_r = 0$ , then the optimal plan is time-consistent, and is given by  $\pi_t = 0, x_t = 0, r_t = r_t^n$  at all times. Thus our arguments for the desirability of low interest-rate volatility are crucial to the argument given here for the desirability of an inertial, non-time-consistent policy. But similar conclusions could be reached even when  $\lambda_r = 0$ , if we assume some other reason for the first-best level of welfare not to be attainable, such as the “cost-push shocks” assumed in Clarida *et al.* (1999).

to the constraints (2.6) and (2.7), so as to minimize  $L_t$ , given the exogenous state  $r_t^n$ , and taking as given the way that the endogenous variables will be chosen (as a function of the economy's exogenous state  $I_T$ ) at all dates  $T > t$ . The evolution of the endogenous variables at later dates is not assumed to depend upon the choices at date  $t$ , because the equilibrium conditions (2.6) and (2.7), that will constrain the later choices, do not involve any lagged endogenous variables, and neither does the period objective function (2.9). The first-order conditions for this optimization problem at each date are given by

$$\pi_t + \phi_{2t} = 0, \quad (3.6)$$

$$\lambda_x(x_t - x^*) + \phi_{1t} - \kappa\phi_{2t} = 0, \quad (3.7)$$

$$\lambda_r(r_t - r^*) + \sigma^{-1}\phi_{1t} = 0, \quad (3.8)$$

instead of (3.2) – (3.4). A Markov-perfect solution is a set of processes  $\{\pi_t, x_t, r_t, \phi_{1t}, \phi_{2t}\}$  that satisfy (2.6), (2.7), and (3.6) – (3.8) at all dates, with the property that each of the endogenous variables just listed depends at date  $t$  *only* upon the part of  $I_t$  that is relevant for forecasting current and future values of the exogenous disturbance  $r_{t+j}^n$ . Note that the absence of the lagged Lagrange multiplier terms from these conditions in *every* period, and not just at date zero, means that there is no problem of time consistency in this case.

The steady-state values of the various variables under the time-consistent plan – *i.e.*, the values they would take in the case that one expected  $r_t^n = 0$  forever – may be obtained by substituting constant values for each variable into the equilibrium conditions and solving. The steady-state values of the endogenous variables are given by

$$r^{tc} = \pi^{tc} = \frac{\lambda_x x^* - \sigma \lambda_r r^*}{\kappa + \lambda_x(1 - \beta)\kappa^{-1} - \lambda_r \sigma}, \quad x^{tc} = \frac{\lambda_x x^* - \sigma \lambda_r r^*}{\lambda_x + (1 - \beta)^{-1}\kappa[\kappa - \lambda_r \sigma]}. \quad (3.9)$$

As is discussed further in the next subsection, these values typically imply a steady-state inflation rate higher than the one that would minimize the steady-state value of the period loss  $L$ .

In determining the equilibrium responses to shocks, it is convenient to work in terms of deviations from steady-state values. Thus we define  $\hat{\pi}_t \equiv \pi_t - \pi^{tc}$ , and so on. The hatted

variables satisfy the same set of difference equations as before, except that we may now drop the constant terms from equations (3.7) – (3.8). If we substitute out the Lagrange multipliers, equations (3.6) – (3.8) imply an equilibrium relation of the form

$$\hat{r}_t = \frac{\kappa}{\lambda_r \sigma} \hat{\pi}_t + \frac{\lambda_x}{\lambda_r \sigma} \hat{x}_t \quad (3.10)$$

among the endogenous variables. Using this to substitute out  $r_t$  in (2.7), we are left with two stochastic difference equations in the endogenous variables  $z_t \equiv [\hat{\pi}_t \ \hat{x}_t]'$ . These may be written in the matrix form

$$E_t z_{t+1} = A z_t + a r_t^n, \quad (3.11)$$

where  $A$  is a  $2 \times 2$  matrix of coefficients, and  $a \equiv [0 \ -\sigma^{-1}]'$ .

Now let the relevant information at date  $t$  about the future evolution of the natural rate be summarized by an exogenous state vector  $s_t$ , with law of motion

$$s_{t+1} = T s_t + \epsilon_{t+1}, \quad (3.12)$$

where  $\epsilon_{t+1}$  is a vector of exogenous disturbances unforecastable at  $t$ , and where the natural rate itself is given by some linear function of these states,

$$r_t^n = k' s_t. \quad (3.13)$$

Then we seek a solution to (3.11) of the form  $z_t = F s_t$ . It follows from (3.11) and (3.12) that the matrix  $F$  must satisfy

$$F T - A F = a k'. \quad (3.14)$$

This is a set of  $2n$  linear equations (where  $n$  is the number of elements in  $s_t$ ) to determine the  $2n$  coefficients  $F$ , and so except in singular cases, a solution exists and is unique.

An instructive simple case is that in which the natural rate follows a first-order autoregressive process,

$$r_{t+1}^n = \rho r_t^n + \epsilon_{t+1}, \quad (3.15)$$

in which case the state vector  $s_t$  consists simply of  $r_t^n$  itself. (We shall throughout the following discussion always assume that  $0 \leq \rho < 1$ .) In this case, there exists a time-consistent plan as long as  $\rho$  is not an eigenvalue of the matrix  $A$ , and it is given by

$$F = -[A - \rho I]^{-1}a.$$

In this solution, the endogenous variables are functions simply of the current value of the natural rate of interest,

$$\hat{\pi}_t = f_\pi r_t^n, \quad \hat{x}_t = f_x r_t^n, \quad \hat{r}_t = f_r r_t^n. \quad (3.16)$$

In the case that

$$\frac{\rho}{(1 - \rho)(1 - \beta\rho)} < \frac{\sigma}{\kappa},$$

*i.e.*, as long as the fluctuations in the natural rate are not *too* persistent, one can show that these coefficients furthermore satisfy

$$f_\pi, f_x > 0, \quad 0 < f_r < 1,$$

regardless of the size of  $\lambda_x$  and  $\lambda_r$ .

This means that interest rates adjust in the direction of the disturbance to the natural rate, but by less than the full amount of the change in the natural rate, owing to the desire to reduce interest-rate variability. (In the limiting case with  $\lambda_r = 0$ , one finds that  $f_r = 1$ , and  $f_\pi = f_x = 0$ .) As a result of the incomplete adjustment of nominal interest rates, an increase in the natural rate stimulates aggregate demand, with the result that both inflation and the output gap increase. Note that in this solution, the only source of persistence in interest-rate fluctuations is the persistence (if any) of the exogenous fluctuations in the natural rate. (Nominal interest rates would be observed to follow a first-order autoregressive process, with autocorrelation coefficient  $\rho$ .) Thus if central bank policy is optimizing in this time-consistent sense, the conventional view described in the introduction is correct – optimizing policy does not involve any inertia in nominal interest rates other than what results from the persistence of the underlying economic conditions in response to which interest rates are properly adjusted.

If one assumes a non-Markovian natural rate process, so that there are additional states in the vector  $s_t$  besides the current natural rate itself, that are relevant for forecasting the future level of the natural rate, then in general the time-consistent solution for  $r_t$  will also depend upon these states. For example, in the case that the matrix  $A$  has both eigenvalues outside the unit circle, the unique bounded solution to (3.11) is obtained by “solving forward” to yield

$$z_t = - \sum_{j=0}^{\infty} A^{-(j+1)} a E_t r_{t+j}^n. \quad (3.17)$$

(Such a solution necessarily exists, and is bounded, as long as the natural rate process itself is bounded.) This expression offers an alternative representation of the time-consistent plan (3.14). Substitution of (3.17) into (3.10) then allows a similar expression

$$\hat{r}_t = -q' \sum_{j=0}^{\infty} A^{-(j+1)} a E_t r_{t+j}^n, \quad (3.18)$$

where  $q'$  is the row vector of coefficients in (3.10), to be derived for the evolution of the nominal interest rate under the time-consistent plan. Expressions (3.17) and (3.18) show clearly that such policy should in general be “forward-looking”, and not simply a function of current conditions alone.<sup>37</sup> At the same time, there is no reason for the central bank’s setting of the nominal interest rate to depend either upon *past* levels of the natural rate (except insofar as these may be variables that help to forecast future levels of the natural rate) or upon past levels of the nominal interest rate itself.

But while this equilibrium could be explained as the result of *optimization* (by a central bank that is unable to commit itself), it does not represent *optimal* policy – for it does not achieve the lowest expected value of the central bank objective function (2.8) consistent with

---

<sup>37</sup>Note, however, that even this representation of optimizing policy gives one no reason to suppose that policy needs to be based upon forecasts of *goal* variables, such as inflation forecasts, as opposed to forecasts of future exogenous disturbances. Note also that the mere fact that one wanted the endogenous variables to respond to information about the future level of the natural rate in the way indicated by (3.17) would not necessarily require explicit response to forecasts *by the central bank*; it could be achieved as a rational expectations *equilibrium* response, owing to forward-looking behavior on the part of the private sector, even if the central bank’s policy rule responds only to current endogenous variables. See the discussion in the next section.

the constraints imposed by the requirements for a rational expectations equilibrium. That *optimal* plan is instead described by the first-order conditions (3.2) – (3.4) derived earlier.

### 3.3 Comparison with the Optimal Non-Inertial Plan

The optimal plan (which requires commitment) is almost invariably different from the time-consistent optimizing plan just characterized, and as we shall see, one respect in which it is different is that it generally involves persistence in the responses of the endogenous variables to shocks that are not due to the persistence of the fluctuations in the natural rate of interest. However, one might ask to what extent the optimal policy is better because it allows for a commitment to inertial behavior, as opposed to being better simply because it allows for *commitment*. For even among the category of plans under which the endogenous variables are functions *only* of the state vector  $s_t$  (in accordance with the dynamic programming principle), commitment generally makes possible a better outcome than that obtained under discretion, as discussed by Clarida *et al.* (1999). Hence it may be worth brief consideration of the optimal *non-inertial* plan.<sup>38</sup>

We shall proceed directly to the case in which the natural rate evolves according to (3.15). In this case, non-inertial plans are those in which each endogenous variable  $y_t$  is a time-invariant linear function of the current natural rate of interest,

$$y_t = y^s + f_y r_t^n, \quad (3.19)$$

where  $y^s$  is the steady-state value of the variable under the optimal “simple” rule, and  $f_y$  indicates its response to fluctuations in the natural rate, as in (3.16). Substituting the representation (3.19) for each of the variables  $y = \pi, x, r$  into (2.6) – (2.7), we find that feasible non-inertial plans correspond to coefficients  $y^s, f_y$  that satisfy

$$(1 - \beta)\pi^s = \kappa x^s, \quad (3.20)$$

---

<sup>38</sup>This is closely related to the question of the optimal “simple” policy rule, considered by Levine (1991). However, there the nature of optimal feedback to the instrument from endogenous variables is considered, whereas we defer this question to section 4.2 below. Furthermore, there the choice of the optimal “simple” rule is allowed to depend upon the state of the economy in the initial period.

$$r^s = \pi^s, \quad (3.21)$$

$$(1 - \beta\rho)f_\pi = \kappa f_x, \quad (3.22)$$

$$(1 - \rho)f_x = -\sigma^{-1}(f_r - 1 - \rho f_\pi). \quad (3.23)$$

Among these plans, we seek the one that minimizes  $E[W]$ , the unconditional expectation of (2.8), taking the unconditional expectation over the stationary distribution of possible initial exogenous states  $r_0^n$ . We take this unconditional expectation so that our choice of the optimal plan does not depend upon the state that the economy happens to be in at the time that the commitment is made.<sup>39</sup>

Given our restriction to non-inertial plans, minimization of  $E[W]$  is equivalent to minimization of  $E[L]$ , the unconditional expectation of the period loss (2.9). Thus we seek to minimize

$$E[L] = [\pi^{s2} + \lambda_x(x^s - x^*)^2 + \lambda_r(r^s - r^*)^2] + [f_\pi^2 + \lambda_x f_x^2 + \lambda_r f_r^2] \text{var}(r^n), \quad (3.24)$$

subject to the linear constraints (3.20) – (3.23). Note that the first term in square brackets in (3.24) involves only the steady-state values  $y^s$ , while the second term involves only the response coefficients  $f_y$ ; similarly, constraints (3.20) – (3.21) involve only the former coefficients, while constraints (3.22) – (3.23) involve only the latter. Thus separate problems define the optimal values of each of the two sets of coefficients.

The optimal non-inertial plan is easily seen to involve a steady state in which

$$r^s = \pi^s = \frac{(1 - \beta)\kappa^{-1}\lambda_x x^* + \lambda_r r^*}{1 + [(1 - \beta)\kappa^{-1}]^2 \lambda_x + \lambda_r},$$

---

<sup>39</sup>If instead we were to minimize  $W$ , conditioning upon the state of the economy at the time of choice as in Levine (1991), the exact non-inertial plan that would be chosen would in general depend upon that state. This is because the choice of how the variables should depend upon  $r^n$  would be distorted by the desire to obtain an initial (unexpected) inflation, without creating expectations of a similar rate of inflation on average in the future; this could be done by exploiting the fact that  $r_0^n$  is known to have a value different from its expected value in the future (which is near zero eventually). By instead defining the optimal non-inertial policy as we do, we obtain a unique policy of this kind, and associated unique values for statistics such as the variability of inflation under this policy. Also, under our definition, unlike Levine's, the optimal "simple" plan is *certainty-equivalent*, just like the fully optimal plan and the time-consistent optimizing plan. That is, the optimal steady-state values  $y^s$  are the same as for a certainty problem, while the optimal response coefficients  $f_y$  are independent of the variance of the disturbance process.

with  $x^s$  then given by (3.20). We observe that in the case that  $\lambda_r$  and/or  $r^*$  are small enough, the steady-state inflation rate satisfies

$$0 < \pi^s < \pi^{tc}.$$

It is slightly positive, as in the calculation of King and Wolman (1996), because (2.6) implies a slightly upward-sloping long-run Phillips curve;<sup>40</sup> but it is smaller than in the discretion equilibrium, because (given  $\beta$  near one) the long-run Phillips curve is much steeper than the short-run trade-off. This indicates that discretionary optimization gives rise to an “inflation bias”, as in the analysis of Barro and Gordon (1983).<sup>41</sup>

Turning to the optimal non-inertial responses to disturbance, the first-order conditions for optimal choice of the  $f_y$  imply that

$$f_r = \frac{\kappa(1 - \beta\rho)^{-1}f_\pi + \lambda_x f_x}{[(1 - \rho)\sigma - \rho\kappa(1 - \beta\rho)^{-1}]\lambda_r}. \quad (3.25)$$

This condition along with (3.22) – (3.23) determines the optimal response coefficients. Note that (3.25) reduces to (3.10) in the case that  $\rho = 0$ , so that the optimal non-inertial policy coincides with the time-consistent policy in this case. However, in general the two solutions do not coincide, and for large  $\rho$  they can be quite different (for example, response coefficients may have opposite signs in the two cases).

---

<sup>40</sup>This is described by (3.20). The effect is quite small, not only because  $\beta$  is plausibly near one, making the long-run Phillips curve steep, but because  $\lambda_x$  is small in a plausible utility-based welfare criterion. The Calvo (1983) model of price-setting is criticized on this ground by McCallum and Nelson (1998), who feel that a plausible model ought to satisfy Friedman’s (1968) “natural rate hypothesis”. The natural rate property is restored in Yun’s (1996) version of the model, that assumes that prices increase at a constant rate between the occasions upon which they are re-calculated, where the constant rate corresponds to the economy’s long-run average rate of inflation. It is interesting to note, however, that even if one feels that the violation of the natural rate hypothesis is reasonable, this provides no ground for preferring a positive average rate of inflation. For, as is shown in the next section, the fully optimal policy implies a steady-state inflation rate that is independent of  $x^*$ .

<sup>41</sup>Note that in the present model, a mere demonstration that  $\pi^{tc} > 0$  does not necessarily imply an inflation bias in this sense, both because positive steady-state inflation raises the steady-state output gap in this model, unlike that of Barro and Gordon, and because we allow for the possibility that  $r^* > 0$ , which also implies advantages to a positive average inflation rate.

### 3.4 Interest-Rate Inertia under the Optimal Plan

The optimal plan corresponds to a bounded solution of the system of difference equations (2.6), (2.7), and (3.2) – (3.4), consistent with the initial conditions (3.5). We begin, as above, by computing the steady-state values of the endogenous variables implied by these equations, by which we mean the constant values that would satisfy the difference equations in all periods (setting aside for the moment the issue of the initial conditions). These steady-state values are given by

$$r^{opt} = \pi^{opt} = \frac{\lambda_r}{\lambda_r + \beta} r^*, \quad x^{opt} = \frac{1 - \beta}{\kappa} \frac{\lambda_r}{\lambda_r + \beta} r^*, \quad (3.26)$$

and the associated steady-state Lagrange multipliers by<sup>42</sup>

$$\phi_1^{opt} = \frac{\beta\sigma\lambda_r}{\lambda_r + \beta} r^*, \quad \phi_2^{opt} = \frac{1 - \beta}{\kappa} \left( \frac{\lambda_x}{\kappa} - \sigma \right) \frac{\lambda_r}{\lambda_r + \beta} r^* - \frac{\lambda_x}{\kappa} x^*.$$

Note that the optimal steady-state inflation rate is independent of  $x^*$ ; this means that as long as  $r^* = 0$ , *i.e.*, the optimal nominal interest rate is the one consistent with zero steady-state inflation, the optimal steady-state inflation rate is *zero*, even when the associated steady-state output level (the natural rate of output) is inefficient. This is true even though a policymaker optimizing in the absence of commitment *would* wish to choose a positive level of inflation (assuming  $x^* > 0$  and  $\lambda_r$  not too large) in order to increase output to a more nearly efficient level; and it is true even though, in the present model (unlike that of Barro and Gordon, 1983), there *is* a long-run Phillips curve trade-off. We obtain the same conclusion if  $\lambda_r = 0$ , and in this case our results agree with those of King and Wolman (1998) for a closely related model with overlapping two-period price commitments. If  $\lambda_r > 0$  and  $r^* \neq 0$ , then  $\pi^{opt} \neq 0$  as well; but this is because of the desirability of a non-zero inflation

---

<sup>42</sup>Note that these values are linear in  $x^*$  and  $r^*$ . Thus, as asserted in the previous section, if the values  $x^*$  and  $r^*$  that measure the inefficiency of the zero-inflation steady state are only of order  $\mathcal{O}(\|\xi\|)$ , then the steady-state values of the Lagrange multipliers are only of order  $\mathcal{O}(\|\xi\|)$  as well. It then follows that in a Taylor series expansion of the exact first-order conditions to characterize the optimal plan, terms of order  $\mathcal{O}(\|\xi\|^2)$  in the structural equations of the model, which contribute terms of order  $\mathcal{O}(\|\xi\|)$  to the partial derivatives of the constraints with respect to the endogenous variables, contribute terms that are only of order  $\mathcal{O}(\|\xi\|^2)$  to the first-order conditions. Hence equations (3.2) – (3.4) represent a first-order Taylor series approximation to the exact conditions, and their solution represents a similar log-linear approximation to the exact optimal plan.

rate in order to allow the nominal interest rate to be nearer its optimal level (say, to reduce the transactions costs associated with economizing on money holdings), and not because of any optimal exploitation of the long-run Phillips curve tradeoff. The contrast between this result and (3.9) indicates the desirability of a central bank's being able to commit itself to a policy that maintains zero average inflation (or at any rate, an average inflation rate that differs from zero only because  $r^* \neq 0$ ), rather than simply acting in the public interest on a period-by-period basis.

One may wonder, of course, about the significance of this “steady-state” solution, given that the equations describing the optimal plan do not generally admit a solution that is constant over time, even when  $r_t^n$  is expected to equal zero at all dates. (This is because the steady-state values of the multipliers are generally inconsistent with the initial conditions (3.5). Only if both  $x^* = 0$  and  $r^* = 0$  do we obtain  $\phi_i^{opt} = 0$  for both  $i = 1, 2$ .) However, in the case of interest to us here – the case in which a bounded optimal plan exists – the optimal plan involves values of the endogenous variables that converge *asymptotically* to the steady-state values, as discussed below. Hence commitment to an optimal plan, in the case that  $r^* = 0$ , means commitment to a rate of inflation that will *eventually* be zero on average.<sup>43</sup> A central bank that wished to behave as *if* it had committed itself to an optimal plan at a date far in the past (as a way of forswearing the temptation to exploit the gains from an unanticipated change in policy) therefore chooses a policy that resulted in zero average inflation, as well.

We turn now to the question of the optimal response to shocks. It will again simplify our equations if we rewrite them in terms of deviations from the steady-state values, defining  $\hat{\pi}_t \equiv \pi_t - \pi^{opt}$ , and so on. The same system of five equations then applies to the hatted variables, except that all constant terms are eliminated from the equations. Using (3.4) to eliminate  $\hat{r}_t$  from (2.7), we are finally left with a system of four stochastic difference equations in the four endogenous variables  $\hat{\pi}_t, \hat{x}_t, \hat{\phi}_{1t}$ , and  $\hat{\phi}_{2t}$ . The system can be written in vector

---

<sup>43</sup>In the case that  $\lambda_r = 0$ , the optimal inflation rate is eventually zero *in all states*, as is argued by King and Wolman (1998). But that is not true here, if we allow a concern for reduction of the variability of interest rates.

form as

$$\begin{bmatrix} E_t z_{t+1} \\ \phi_t \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} z_t \\ \phi_{t-1} \end{bmatrix} + \begin{bmatrix} a \\ 0 \end{bmatrix} r_t^n, \quad (3.27)$$

where  $\phi_t \equiv [\hat{\phi}_{1t} \ \hat{\phi}_{2t}]'$  is the vector of deviations of the values of the Lagrange multipliers,  $A$  and  $a$  are the same matrix and column vector as in (3.11) above, and  $B, C$ , and  $D$  are other  $2 \times 2$  matrices of coefficients. Here the bottom two equations are equations (3.3) and (3.2), in that order, expressed as laws of motion for the Lagrange multipliers. The top two equations are again (2.6) and (2.7), in that order, expressed as expectational difference equations for inflation and the output gap. The first equation is substituted into the second to eliminate  $E_t \hat{\pi}_{t+1}$ , and the third equation is similarly substituted into it to eliminate  $\hat{\phi}_{1t}$ . The non-zero elements of  $B$  (which occur only in the second row) result from this last substitution.

Given a bounded stochastic process for the exogenous disturbance  $r_t^n$ , one can show using the methods of Blanchard and Kahn (1980) that the system (3.27) has a unique bounded solution if and only if the matrix

$$M \equiv \begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

has exactly two eigenvalues outside the unit circle, and two that are inside.<sup>44</sup> (This is because the state vector contains exactly two predetermined variables, the two elements of  $\phi_{t-1}$ .) One can show that the matrix  $M$  necessarily has two eigenvalues with modulus greater than  $\beta^{-1/2}$  and two with modulus smaller than this. This means that there are necessarily *at least* two eigenvalues outside the unit circle, and so there is *at most* a single bounded solution. (This is as we should expect, since any bounded solution would have to correspond to an optimal plan.) The conditions required for there to be only two eigenvalues outside the unit circle, so that a bounded solution does exist, seem likely to be satisfied for reasonable parameter values,<sup>45</sup> and since our Taylor series approximation methods can only validly be employed in this case, we assume that they are satisfied.

---

<sup>44</sup>Here we neglect to discuss certain non-generic cases that can occur only for special parameter values.

<sup>45</sup>For example, they are satisfied by the “calibrated” parameter values used for the numerical computations below.

In this case, the unique bounded solution can be written in the form

$$z_t = G\phi_{t-1} - \sum_{j=0}^{\infty} \tilde{A}^{-(j+1)} a E_t r_{t+j}^n, \quad (3.28)$$

where  $G$  and  $\tilde{A}$  are  $2 \times 2$  matrices. Furthermore, the eigenvalues of the matrix  $\tilde{A}$  are exactly the two eigenvalues of  $M$  that lie outside the unit circle, so that the infinite sum converges in the case of any bounded process for the natural rate. Substituting this solution for  $z_t$  into the last two equations of (3.27), we obtain a law of motion for the Lagrange multipliers of the form

$$\phi_t = N\phi_{t-1} - C \sum_{j=0}^{\infty} \tilde{A}^{-(j+1)} a E_t r_{t+j}^n, \quad (3.29)$$

where  $N \equiv CG + D$  is a  $2 \times 2$  matrix the eigenvalues of which are exactly the two eigenvalues of  $M$  that lie inside the unit circle. This last property of the matrix  $N$  implies that (3.29) defines a bounded stochastic process for the multipliers  $\phi_t$ , given any bounded process for the natural rate. Then (3.29) describes the evolution of the multipliers, given the initial condition (3.5) and an exogenous process for the natural rate, while (3.28) describes the evolution of the endogenous variables  $z_t$ , given the evolution of the multipliers and an exogenous process for the natural rate. Given this evolution of the variables  $z_t$  and  $\phi_t$ , the implied evolution of the central bank's instrument  $r_t$  is easily derived, and can also be written in the form

$$\hat{r}_t = p'\phi_{t-1} - q' \sum_{j=0}^{\infty} \tilde{A}^{-(j+1)} a E_t r_{t+j}^n. \quad (3.30)$$

where  $q'$  is the same vector of coefficients as in (3.18), and  $p'$  is another vector of two coefficients.<sup>46</sup>

Note that the second term on the right-hand side of (3.28) is of a similar form as (3.17), and likewise for the second term on the right-hand side of (3.30) and (3.18). (The only difference in these terms is the replacement of the matrix  $A$  by  $\tilde{A}$ .) The crucial *qualitative* difference is the presence of the  $\phi_{t-1}$  terms as well in the case of the optimal plan. These terms imply that the endogenous variables at date  $t$  – and in particular, the central bank's setting

---

<sup>46</sup>It is easily seen from (3.4) that  $p'$  must equal  $-(\lambda_r \sigma)^{-1}$  times the first row of the matrix  $N$ . In fact,  $q'$  is the same multiple of the first row of  $C$ .

of the interest rate at that date – should *not* depend solely upon current and forecasted future values of the natural rate of interest. They should also depend upon the predetermined state variables  $\phi_{t-1}$ , which represent an additional source of inertia in optimal monetary policy, independent of any inertia that may be present in the exogenous disturbance process  $r_t^n$ . The additional terms represent the way in which policy should deviate from what would be judged optimal simply taking into account the *current* outlook for the economy, in order to follow through upon *commitments* made at an earlier date. It is the desirability of the central bank's being *able* to credibly commit itself in this way that makes it desirable for monetary policy to be somewhat inertial.

The extent to which these equations imply inertial behavior of the nominal interest rate can be clarified by writing a law of motion for the interest rate that makes no reference to the Lagrange multipliers. Let us assume again a state-space representation (3.12) – (3.13) for the evolution of the natural rate. Equation (3.29) then takes the form

$$\phi_t = N\phi_{t-1} + ns_t, \quad (3.31)$$

where the matrix of coefficients  $n$  is given by

$$n \equiv -C \sum_{j=0}^{\infty} \tilde{A}^{-(j+1)} ak' T^j.$$

The endogenous variable  $\phi_{2t}$  may be eliminated from the system of equations (3.31), yielding an equation with instead two lags of  $\phi_{1t}$ . We may write this last equation as

$$\det[I - NL]\phi_{1t} = n'_1 s_t + (N_{12}n'_2 - N_{22}n'_1)s_{t-1},$$

where  $L$  is the lag operator, and  $n'_i$  is the  $i$ th row of the matrix  $n$ . Then using (3.4) to substitute out  $\phi_{1t}$ , we obtain the law of motion

$$Q(L)\hat{r}_t = R(L)s_t \quad (3.32)$$

for the nominal interest rate, where

$$Q(L) \equiv \det[I - NL], \quad R(L) \equiv -(\lambda_r \sigma)^{-1} [n'_1 + (N_{12}n'_2 - N_{22}n'_1)L]. \quad (3.33)$$

This should be compared with the simple result (3.13) in the case of the time-consistent plan.

Equation (3.32) implies that under the optimal plan, there are *intrinsic* dynamics to the evolution of the nominal interest rate, unrelated to any persistence in the fluctuations in the exogenous states  $s_t$ . For example, in the case that  $r_t^n$  is an exogenous white-noise process (unforecastable at any prior date), (3.32) implies that  $r_t$  should instead follow a second-order autoregressive process. The degree of persistence of these intrinsic dynamics is determined by the roots  $\mu_i$  of the characteristic equation

$$Q(\mu) = 0$$

associated with the autoregressive polynomial in (3.32). These roots are just the eigenvalues of the matrix  $N$ , or the two eigenvalues of  $M$  that have a modulus less than one. The implied degree of monetary policy inertia is greater the larger are these roots. These roots are determined by factors independent of the dynamics of the exogenous disturbances. Thus it may be optimal for nominal interest rates to exhibit a great deal of persistence, regardless of the degree of persistence of the fluctuations in the natural rate.

### 3.5 A Simple Limiting Case

The extent to which the equations just derived imply behavior that might appear to involve interest-rate “smoothing” can be clarified by considering a limiting case, in which a simple closed-form solution may be obtained. This is the limiting case in which the value of the parameter  $\kappa$  (the slope of the “short-run Phillips curve”) approaches zero. In this limit, variations in output relative to potential cause no change in the level of real marginal cost, and firms accordingly have no reason to change their prices at any time. Hence  $\pi_t = 0$  at all times, regardless of monetary policy. We shall assume that the values of all other parameters are unchanged.<sup>47</sup>

---

<sup>47</sup>This does not necessarily make sense, if the coefficients in the loss function (2.9) are intended to represent an approximation to true social welfare, as suggested above, since in that case there is a theoretical relation between  $\lambda_x$  and the various preference and technology parameters that also determine  $\kappa$ . But even in that

In this limiting case, the  $\kappa\phi_{2t}$  term in (3.3) can be neglected, so that it becomes possible to solve for the variables  $x_t, r_t$ , and  $\phi_{1t}$  using only equations (2.7), (3.3), and (3.4).<sup>48</sup> This system of equations can again be written in the form (3.27), but now  $z_t$  is simply the scalar variable  $\hat{x}_t$  and  $\phi_t$  is simply the scalar variable  $\hat{\phi}_{1t}$ . As a consequence,  $A, B, C, D$ , and  $a$  are all now scalars, given by

$$\begin{aligned} A &= 1 + \lambda_x(\lambda_r\sigma^2)^{-1}, \\ B &= -(\beta\lambda_r\sigma^2)^{-1}, \\ C &= -\lambda_x, \\ D &= \beta^{-1}, \\ a &= -\sigma^{-1}. \end{aligned}$$

The characteristic equation for  $M$  is then simply

$$\mu^2 - [1 + \beta^{-1} + \lambda_x(\lambda_r\sigma^2)^{-1}]\mu + \beta^{-1} = 0. \quad (3.34)$$

One observes that it necessarily has two real roots, satisfying

$$0 < \mu_1 < 1 < \beta^{-1} < \mu_2,$$

and that  $\mu_2 = (\beta\mu_1)^{-1}$ . Because exactly one root is inside the unit circle, the system of equations (3.27) has a unique bounded solution. The solution is given by equations (3.28) and (3.29), where now  $G, N$ , and  $\tilde{A}$  are scalars

$$\begin{aligned} G &= (\beta^{-1} - \mu_1)\lambda_x^{-1} > 0, \\ N &= \mu_1 > 0, \\ \tilde{A} &= \mu_2 > 1. \end{aligned}$$

---

case, we may view the present calculation as solving a mathematical problem that is closely similar in form, even if not identical, to one that arises in an optimal monetary policy problem.

<sup>48</sup>The intuition for this reduction in the order of the system of equations is simple. We no longer need to solve for  $\pi_t$ , as we set this variable equal to zero. We no longer obtain a first-order condition corresponding to (3.2) by differentiating the Lagrangian with respect to  $\pi_t$ , because it is not possible to vary inflation. Likewise, we no longer obtain a Lagrange multiplier  $\phi_{2t}$  corresponding to constraint (2.6), as this constraint is already guaranteed to be satisfied once we have set  $\pi_t$  equal to zero at all times.

Since equation (3.29) now involves only  $\hat{\phi}_{1t}$ , it is possible to use (3.4) to substitute  $\hat{r}_t$  for  $\hat{\phi}_{1t}$ , and thus directly obtain an equation for the optimal interest-rate dynamics,

$$\hat{r}_t = \mu_1 \hat{r}_{t-1} + \lambda_x (\lambda_r \sigma^2)^{-1} \sum_{j=0}^{\infty} \mu_2^{-(j+1)} E_t r_{t+j}^n. \quad (3.35)$$

This gives us a law of motion of the form (3.32), but in this limiting case, a representation is possible in which  $Q(L)$  is only of first order, and  $R(L)$  is a constant (there are no lags at all). In fact, one can easily show that (3.35) is a partial-adjustment equation of the form (1.2), where the inertia coefficient  $\theta = \mu_1$ , and the time-varying interest-rate “target” is given by<sup>49</sup>

$$\bar{r}_t = r^{opt} + (1 - \mu_2^{-1}) \sum_{j=0}^{\infty} \mu_2^{-j} E_t r_{t+j}^n. \quad (3.36)$$

Thus the optimal interest-rate dynamics are described by partial adjustment toward a moving average of current and expected future natural rates of interest.

In the case that the natural-rate dynamics are of the simple form (3.15), the target rate is just a function of the current natural rate of interest, although (because of expected mean-reversion of the natural rate in the future) it varies less than does the natural rate itself. Specifically, we have

$$\bar{r}_t = r^{opt} + k r_t^n, \quad (3.37)$$

where  $k \equiv (\mu_2 - 1)/(\mu_2 - \rho)$ , so that  $0 < k < 1$ . If the fluctuations in the natural rate are largely transitory, the elasticity  $k$  may be quite small, though it is any event necessarily greater than  $1 - \beta$ . If the fluctuations in the natural rate are nearly a random walk ( $\rho$  is near one), the elasticity  $k$  instead approaches one. In this case, interest rates eventually change by nearly as much as the (nearly permanent) change that has occurred in the natural rate; but even in this case, the change in the level of nominal interest rates is delayed. As a result, an innovation in the natural rate is followed by a series of interest rate changes in the same direction, as in the characterizations of actual central-bank behavior by Rudebusch and Goodhart.

---

<sup>49</sup>This representation is possible because (3.34) implies that  $\lambda_x (\lambda_r \sigma^2)^{-1}$  is equal to  $(1 - \mu_1)(\mu_2 - 1)$ .

While this partial-adjustment representation is only exactly correct in an unrealistic limiting case, it provides considerable insight into the optimal interest-rate responses in more realistic cases. This can be shown through numerical analysis of a case with  $\kappa > 0$ .

### 3.6 A Numerical Example

To consider what degree of interest-rate inertia might be optimal in practice, it is useful to consider a numerical example, “calibrated” to match certain quantitative features of the Rotemberg and Woodford (1997, 1998) analysis of optimal monetary policy for the U.S. economy. The numerical values that we shall use are given in Table 1. The value of  $\beta$  must be only slightly less than one, given observation of only a small positive average real rate of return. The values for  $\sigma$  and  $\kappa$  represent the estimates of Rotemberg and Woodford, who find that these values result in the best fit between the estimated impulse responses of inflation and output to a monetary policy (identified using structural VAR methodology) and those predicted by their model (which, insofar as the role of these two parameters is concerned, is essentially the same as that considered here).

The assumed standard deviation of fluctuations in the natural rate follows from the Rotemberg-Woodford estimates of the statistical properties of the residuals of their structural equations.<sup>50</sup> The value of this parameter has in any event no effect upon our conclusions about the *relative* variability of different variables under alternative policies, since there is only one kind of stochastic disturbance.<sup>51</sup>

More important are our assumptions about the serial correlation properties of the shocks. We shall assume an AR(1) process as in (3.15), as a result of which we need only calibrate

---

<sup>50</sup>Specifically, this number represents their estimate of the standard deviation of  $E_{t-2}r_t^n$ , where  $r_t^n$  is defined as in (3.13). This is because their structural equations coincide with those of the simpler model used here only when conditioned upon information available two quarters earlier. Because of the two-period lag assumed in the effects of monetary policy upon output and inflation in their variant model,  $E_{t-2}r_t^n$  is the exogenous disturbance process in their model that plays the role most analogous to that of the “natural rate of interest” in the present model.

<sup>51</sup>However, our assumption about the variability of the natural rate plays an important role in justifying our assumed value for  $\lambda_r$ . The value assumed here implies that the standard deviation of the natural rate is roughly the same size as its mean. Hence it is frequently negative, meaning that in a zero-inflation environment the zero bound on nominal interest rates would prevent short rates from perfectly tracking the natural rate of interest.

a single parameter  $\rho$ . The value of most interest here is unclear, since in reality there are probably different types of disturbances, with differing degrees of persistence.<sup>52</sup> We shall accordingly compare our results under different assumed values for  $\rho$ . Our baseline value of  $\rho = .35$ , however, is chosen so as to imply, in the simpler model used here, a degree of concern for reduction of interest-rate variability similar to that obtained by Rotemberg and Woodford in their estimated model. We have sought a value for  $\rho$  such that optimal monetary policy (i) involves a similar degree of interest-rate variability as in their numerical results, and (ii) implies a similar shadow value of relaxation of the interest-rate variability constraint (2.11). The value  $\rho = .35$  is reasonably satisfactory on both grounds.<sup>53</sup> Below, however, we also report numerical results for the values  $\rho = 0$  and  $\rho = .9$ . These values are chosen to illustrate how our conclusions depend upon the assumed degree of serial correlation of the shocks.

The coefficient  $\lambda_x$  is the theoretical relative weight on output gap variability obtained by Taylor series expansion of the expected utility of the representative household, evaluated using the structural parameters of their model; it corresponds to a coefficient from a similar Taylor series expansion reported in the Rotemberg-Woodford welfare analysis. And finally,  $\lambda_r$  is the Lagrange multiplier associated with interest-rate variability, if we seek to minimize the expected discounted sum of welfare losses (2.10) subject to the constraint (2.11), as discussed in section 2.<sup>54</sup> We impose this constraint assuming the value of  $K = 1.44$  as discussed above, and assuming the values given in Table 1 for the other parameters. Hence, in our baseline case ( $\rho = .35$ ), the policy that minimizes our welfare criterion  $E[W]$  also

---

<sup>52</sup>The shock process estimated by Rotemberg and Woodford has two innovations per period and complicated dynamics. See Appendix 2 of their NBER working paper.

<sup>53</sup>As shown in Table 2, this value implies that under the optimal policy,  $V[r] = 1.921$ , whereas Rotemberg and Woodford obtain  $var(r) = 1.928$  for their optimal policy. (The measure of interest-rate variability  $V[r]$  used in this paper is defined below in (3.38).) A slightly higher value of  $\rho$ , but in any event well below .4, would match this precisely. At the same time, this value implies a shadow value (in terms of increasing  $E[\tilde{W}]$ ) of increased interest-rate variability  $V[r]$  equal to .236, as reported in Table 1. This compares with the value .224 reported by Rotemberg and Woodford. A slightly lower value of  $\rho$ , but in any event well above .3, would match this precisely.

<sup>54</sup>To be precise, this measures the marginal increase in  $E[\tilde{L}]$  that is possible per unit marginal increase in  $V[r]$ , the variability measure defined below in (3.38). The unconditional expectation that is taken of the loss measure is the same as that discussed below for the statistic  $E[W]$  reported in Table 2.

minimizes the expected discounted sum of (2.10), our quadratic approximation to the utility of the representative household, subject to the constraint implied by the zero bound on nominal interest rates. We could similarly calibrate the parameters  $x^*$  and  $r^*$ , but as these do not matter for the optimal responses to shocks, we omit discussion of them.

For the parameter values in Table 1, the matrix  $N$  is given by

$$N = \begin{bmatrix} .4611 & .0007 \\ -.7743 & .6538 \end{bmatrix},$$

and its eigenvalues are found to be approximately .65 and .46. Both of these are substantial positive quantities, suggesting that once interest rates are perturbed in response to some shock, it should take several quarters for them to be restored to nearly their normal level, even if the shock is completely transitory.

Figure 1 illustrates this by showing the optimal response of the short-term nominal interest rate to a purely transitory increase in the natural rate of interest. This might be due either to a temporary increase in the autonomous component of spending,  $\hat{G}_t$ , or to a temporary decrease in the natural rate of output due to an adverse “supply shock”. Either type of shock would imply a temporary increase in the equilibrium real rate of interest in a flexible-price model, and, in our model with sticky prices, will increase both the output gap and inflation,<sup>55</sup> in the absence of an offsetting adjustment of monetary policy.

To be precise, the Figure displays the impulse response of  $\hat{r}_{t+j}$ , for  $j = 0$  through 10, to a unit positive innovation in  $\epsilon_t$ , where the law of motion for the natural rate is given by (3.15) with autocorrelation coefficient  $\rho = 0$ . The shock in question unexpectedly raises the natural rate  $r_t^n$  by one percentage point, but the natural rate is expected to return to its normal level by the next quarter, as shown by the dashed impulse response. The disturbance to the nominal interest rate is equally transitory under the optimal non-inertial plan, also shown in the figure by a dash-dotted line, though the amplitude of the nominal interest rate increase would be smaller than the increase in the natural rate, in order to reduce interest-rate variability at the price of some increase in inflation and output gap variability. The

---

<sup>55</sup>Note that an adverse supply shock *lowers* output, but increases the output *gap*, as sticky prices prevent output from falling as much as it would in the case of fully flexible prices.

impulse response under an optimal policy is instead given by the solid line. Nominal interest rates rise immediately, but by only a fraction of the increase in the natural rate (about 24 basis points). However, the optimal path of nominal interest rates involves a more *persistent* increase than in the case of the natural rate. While the natural rate is perturbed only in the quarter of the shock, the central bank keeps nominal interest rates 11 basis points above their normal level in the following quarter, and still 5 basis points above normal even two quarters later.

The associated effects of the shock upon inflation and the output gap are shown by the solid impulse responses in Figure 2, which also reproduces the impulse response of the nominal interest rate in the third panel.<sup>56</sup> In the quarter of the shock, the failure of interest rates to rise as much as the increase in the natural rate results in a temporarily negative “interest-rate gap”, but by the next quarter the interest-rate gap has become positive (interest rates higher than the natural rate), owing to the inertia in central bank policy. The fact that a negative interest-rate gap is allowed to develop in the quarter of the shock results in a temporary positive output gap; but in the next quarter and after, the positive (current and expected future) interest-rate gaps result in *negative* output gaps. The resulting response of inflation involves only a very small increase in the quarter of the shock, and then a sharper decrease in the inflation rate after the positive interest-rate gap develops. In the quarter of the shock, the inflationary effect of the positive output gap is offset by the disinflationary effect of the *anticipation* of negative output gaps in the later quarters. It is this effect of the private sector’s anticipation of inertial monetary policy that accounts for the desirability of such inertia – for it allows the central bank to offset the inflationary impact of the shock *without* having to raise nominal interest rates by nearly as much as the increase in the natural rate, as it would have to in order to prevent increased inflation in the quarter of the shock through an interest-rate increase that is expected to be no more persistent than the shock.

---

<sup>56</sup>Here the inflation rate and nominal interest rate are reported as annualized percentage rates; thus the responses plotted are actually for  $4\hat{\pi}_t$  and  $4\hat{r}_t$ .

To further clarify the advantages of interest-rate inertia, Figure 2 also plots the impulse responses associated with the optimal non-inertial policy, assuming the same parameter values. (These impulse responses are plotted as dash-dotted lines in each panel.) Note that the responses of all variables to the shock are purely transitory under the non-inertial plan, as they would be under the time-consistent plan. (Note that the optimal non-inertial plan coincides with the time-consistent plan, in the case that  $\rho = 0$ .) The shock would result in a much larger increase in inflation in the quarter of the shock than occurs under the optimal policy, even though the central bank also raises nominal interest rates more sharply under this plan. This shows the cost to the central bank of not being able to commit itself to an inertial policy. The initial positive output gap is larger under the non-inertial plan as well. Of course, the non-inertial plan has the advantage that the disturbances to inflation, the output gap, and the nominal interest rate are purely transitory. Still, the overall variability of inflation is lower under the optimal plan than under the non-inertial one; for even though inflation falls *below* its normal level in the quarter following the shock, under the optimal plan, the decline in this quarter under the optimal plan is not as large as the increase in inflation during the quarter of the shock under the time-consistent plan.

Statistics regarding the variability of the various series under the two plans are reported in the first panel of Table 2. Here independent drawings from the same distribution of shocks  $\epsilon_t$  are assumed to occur each period,<sup>57</sup> and infinite-horizon stochastic equilibria are computed under each policy. The measure of variability reported for each variable  $z_t$  is

$$V[z] \equiv E[E_0\{(1 - \beta) \sum_{t=0}^{\infty} \beta^t z_t^2\}], \quad (3.38)$$

where the outer (unconditional) expectation is over possible initial states of the economy  $r_0^n$  at the time that policy is chosen, computed using the stationary distribution associated with the exogenous process (3.15) for the natural rate. The unconditional expectation allows us a measure that is independent of the economy's initial state. Except for the discounting,  $E[z]$

---

<sup>57</sup>The assumed variance of the shocks  $\epsilon_t$  is chosen so as to imply a standard deviation for the natural rate of interest of the size indicated in Table 1. This is irrelevant, of course, for the comparisons of *relative* variances between the two regimes, or comparison of the relative size of expected losses  $E[W]$ .

corresponds to the unconditional variance of  $z_t$ , and in the case of non-inertial plans, it is equal to the unconditional variance even though  $\beta < 1$ . In the case of the optimal plan, the discounted measure is of greater interest, because our loss measure  $E[W]$  is in that case just a weighted sum of the previous three columns.

Here  $E[W]$  is again the unconditional expectation of (2.8), integrating over the stationary distribution for the initial state  $r_0^n$ . Since the optimal plan minimizes  $W$  regardless of the initial state  $r_0^n$ , it minimizes  $E[W]$  among all possible plans consistent with the structural equations. The optimal non-inertial plan also minimizes  $E[W]$  among all non-inertial plans; thus there is a necessary ranking of this loss measure, among the three plans for which statistics are reported for each value of  $\rho$ . The values reported in Table 2 are computed for the case in which  $x^* = r^* = 0$ , so that the steady-state values of the variables are the same under all three plans. Thus the reported statistic  $E[W]$  refers only to losses associated with the economy's response to transitory shocks under each of the three policies. Alternatively, the final column reports the amount by which  $E[W]$  exceeds the steady-state value (the value of  $W$  achieved if  $r_t^n = 0$  with certainty forever). This is the statistic of interest if one wants to compare fully optimal policy under commitment to a regime of discretionary optimization (for example) in which the values of  $x^*$  and  $r^*$  have been adjusted to make the steady state under discretion coincide with the optimal steady state. (Were we not to make this assumption, the welfare losses associated with time-consistent and optimal non-inertial policies would exceed those associated with the optimal policy by an extent even larger than is indicated in Table 2, owing to the distortions of the steady state discussed above.)

One observes that the variability of inflation is nearly 75 percent larger under the time-consistent policy, or the optimal non-inertial policy, than under the optimal policy, while the variance of the nominal interest rate is more than twice as large. (The variance of the output gap is more than a third larger as well, although, since the variability of this variable has only a small weight in our assumed loss function, the optimal policy is not primarily chosen with a view to stabilization of the output gap.) Hence the loss measure  $E[W]$  is 60 percent higher under the time-consistent or non-inertial policy.

Figures 3 and 4 present the corresponding impulse response functions in the case of a natural rate process that exhibits positive serial correlation, specifically, our baseline case described by (3.15) with an autoregressive parameter  $\rho = .35$ . In this case the advantages of optimal policy over the time-consistent plan are more pronounced, though commitment to an optimal non-inertial policy can also do somewhat better in this case than the discretionary outcome. Again one observes that under an optimal policy regime, the nominal interest rate does not at first rise as much as the increase in the natural rate; then, as the natural rate declines back toward its normal level, the central bank continues to keep nominal interest rates high for a time, so that they return to normal more slowly than does the natural rate. Thus optimal policy results in a negative interest-rate gap at first, followed by a positive interest-rate gap later. Again, this is associated with an initial positive output gap, followed by a smaller negative output gap a few quarters later, and a very small initial increase in inflation, followed by a subsequent decrease in inflation to a level lower than its steady-state level.

And once again, this pattern results in lower overall variability of inflation, the output gap, *and* of the nominal interest rate, than one would obtain under the time-consistent optimizing policy. In fact, the relative advantage of the optimal policy is even greater in this case: the variance of inflation is nearly four times as large under the time-consistent plan, and the variances of both the output gap and the nominal interest rate are more than twice as large. The result is that the central bank's loss function is two-and-a-third times as large under the time-consistent policy.

The picture is somewhat different if the central bank commits itself to an optimal non-inertial policy. In this case, it should commit itself to raise interest rates more in response to an increase in the natural rate than occurs under the time-consistent plan; the result is that the responses of the interest-rate gap, inflation, and the output gap are all smaller in amplitude by a factor of about two-thirds. The result is substantially less variability of inflation and the output gap, but even more variability of interest rates, achieving a small, but rather modest, reduction in  $E[W]$ . The optimal plan is still significantly better in terms

of both inflation variability and interest-rate variability (though output gap variability, which matters relatively little for welfare under our parameter values, is slightly higher), and the overall loss measure  $E[W]$  is still more than twice as large under the non-inertial policy. These comparisons clearly show the advantages, from the point of view of improved stabilization, not only of the central bank's being able to commit itself, but of its committing itself to an inertial policy.

Finally, Figures 5 and 6 present the corresponding impulse response functions in the case of an even more persistent natural rate process, with  $\rho = .9$ . In the case of these very persistent fluctuations in the natural rate of interest, time-consistent optimizing behavior does extremely poorly at stabilizing the central bank's goal variables. The variance of inflation is a thousand times as large as under the optimal plan, and the variance of the interest rate is over 60 times as large; as a result, the expected value of the loss criterion is more than 230 times as large.<sup>58</sup>

The dramatic failure of discretionary optimization in this case occurs because, for these parameter values, the time-consistent optimizing response actually involves an interest-rate increase that is *larger* than the increase in the natural rate, despite the positive weight  $\lambda_r$  in the central bank's loss function. This occurs because of a "discretion trap" similar to the one analyzed by Barro and Gordon (1983), despite the fact that it involves no incorrect target level of output or bias in the average rate of inflation. When an inflationary real disturbance occurs, the expectation that the central bank will allow high inflation in *subsequent* quarters (while the natural rate of interest continues to be high) creates a situation in which the central bank feels compelled to choose high inflation in the present; given the inflationary expectations, the only policy that could achieve lower inflation would involve an unacceptably large increase in interest rates. But the expectation that the central bank will behave this way results in an equilibrium in which the central bank must accept very high (nominal)

---

<sup>58</sup>Given the highly volatile behavior of the endogenous variables under this solution, one may doubt the accuracy of our linear-quadratic approximations. Still, the qualitative conclusion is surely correct; for if there were a solution with low inflation and interest-rate volatility, then our approximate methods *would* accurately describe it.

interest rates, as a result of the persistent high inflation.

Interestingly, in the case of a shock as persistent as this, the optimal plan does not involve nominal interest rates that ever rise to the same degree as does the natural rate. However, the optimal response does again involve deflation after the initial quarter, brought about by a negative output gap that is expected to be maintained from late in the first year onward. This low level of output is achieved because short-term *real* interest rates are high (rising by 70 basis points during the first year after the shock). Real interest rates are high, of course, because expected inflation is low, rather than because nominal interest rates are high (by comparison with what happens under the time-consistent plan). This illustrates that it is not simply the path of nominal interest rates *alone* that distinguishes optimal policy, a topic to which we return in the next section.

The instability here of the time-consistent plan is mainly an indication that discretionary optimization is severely suboptimal, rather than an indication that a non-inertial plan must be extremely unstable. Indeed, the optimal non-inertial plan is not too much worse than the fully optimal plan, and is worse only insofar as it involves higher interest-rate volatility. As Figure 6 shows, the optimal non-inertial plan involves a commitment to *lower* inflation and a *negative* output gap when the natural rate is above average, results that are achieved by raising *real* interest rates by even more than the increase in the natural rate. Such a response, however, cannot be achieved endogenously by a non-inertial “Taylor rule” that raises nominal interest rates only insofar as either inflation or the output gap increase. The “Taylor rule” would have to be augmented by a term that is increasing in the current natural rate of interest, with a coefficient *greater* than one. Hence achievement of these stabilization gains through commitment to a non-inertial policy rule would depend critically upon accurate observation of the current exogenous shock. But as we show below, the optimal plan can be implemented by an inertial feedback rule that requires no feedback at all from the exogenous shock, and this may be another important advantage of inertial policy.

Visual inspection of the optimal impulse responses in Figures 1, 3 and 5 suggests that “partial adjustment” of the nominal adjustment toward a level determined by the current

natural rate of interest, just as in the limiting case analyzed above, gives a reasonable approximate account of optimal interest-rate dynamics. The reason for this is not hard to see from our previous analysis. Note that the element  $N_{12}$  of the matrix  $N$  is quite small; it is three orders of magnitude smaller than the other elements of the matrix. Hence one would not obtain too poor an approximation to the optimal interest-rate dynamics by setting  $N_{12}$  equal to zero in expression (3.32). But in that case,  $Q(L)$  and  $R(L)$  reduce to

$$Q(L) = (1 - N_{11}L)(1 - N_{22}L),$$

$$R(L) = -(\lambda_r\sigma)^{-1}n'_1(1 - N_{22}L),$$

as a result of which (3.32) is equivalent to

$$\hat{r}_t = N_{11}\hat{r}_{t-1} - (\lambda_r\sigma)^{-1}n'_1s_t. \quad (3.39)$$

This implies partial adjustment of the interest rate, as in equation (1.2), toward a time-varying “target” interest rate equal to

$$\bar{r}_t = r^{opt} - [(1 - N_{11})\lambda_r\sigma]^{-1}n'_1s_t, \quad (3.40)$$

with an inertia coefficient of  $\theta = N_{11}$ . In the case that the natural rate evolves according to (3.15), this time-varying target rate is again described by an equation of the form (3.37). In our numerical example, in the case that  $\rho = 0$ , the target rate (neglecting the constant term) is given by  $\bar{r}_t = .44r_t^n$ , while in the case that  $\rho = .35$ , the elasticity is instead  $k = .52$ , and when  $\rho = .9$ ,  $k = .75$ . In each case, the inertia coefficient is equal to  $\theta = .46$ , indicating that interest rates should be adjusted only about half of the way toward the current target level (implied by the natural rate) within the quarter.

## 4 Advantages of a Central Bank Preference for Smoothing

We turn now to the question of how the optimal responses to shocks, derived in the previous section, are to be brought about in practice. There are various ways in which this question

of implementation might be approached. Perhaps the most straightforward is the design of a policy *rule* to which the central bank may commit itself; this approach is treated in the next section. But another approach is to conceive the choice of monetary policy as a delegation problem. That is, one asks what *sort* of central banker (or monetary policy committee) should be appointed, taking it as given that the central banker will seek to maximize the good as he or she personally conceives it; or one asks what *goal* the central bank should be charged with, understanding that the details of the pursuit of the goal on a day-to-day basis should then be left to the bank. This sort of question is of considerable practical relevance, especially since the central banks that are currently most explicit about their commitment to systematic monetary policy (such as the “inflation targeting” central banks (Svensson (1999)), generally make commitments about goals rather than the details of their operating procedures.

The optimal delegation question becomes non-trivial if one assumes, as is common in this literature, that the central bank will pursue its goal in a *discretionary* fashion, rather than committing itself to an optimal plan, so that the outcome for the economy will be the time-consistent plan associated with that goal.<sup>59</sup> In this case, the optimal goal with which to charge the central bank need not correspond to the true social welfare function; inefficient (discretionary) pursuit of a *distorted* objective may produce a better outcome, from the standpoint of the true social objective, than inefficient pursuit of the true objective itself.<sup>60</sup> For example, Rogoff (1985) famously argues that appointment of a central banker who assigns a greater relative weight to inflation stabilization than does the true social welfare function may actually better serve social welfare, as an appropriate degree of bias can help to offset the inflationary bias due to discretion.<sup>61</sup> Svensson (1996) similarly argues

---

<sup>59</sup>It is, however, sometimes disputed whether actual central bankers are as little capable of choosing to commit themselves as such models assume; see, e.g., McCallum (1999, sec. 2).

<sup>60</sup>This somewhat paradoxical conclusion is familiar in the theory of delegation more generally. See, e.g., Persson and Tabellini (1994).

<sup>61</sup>Walsh’s (1995) contracting solution to the problem of inflation bias under discretion also amounts to arranging for the central banker to pursue an objective other than social welfare. In that formulation, the state-contingent payments to the central banker serve to modify the bank’s objective function in the desired direction. Here we shall speak of choosing the central bank’s objective, without specifying how this is to be done. It might be through a contract, through screening of the personalities of central bankers before

that a central bank with a price level target, rather than an inflation target, may better serve to stabilize inflation in response to shocks, owing to the bias resulting from discretion in the central bank's response to shocks.<sup>62</sup>

We have shown above that in the case of the policy problem considered here, discretionary pursuit of the true social objective (2.8) similarly results in a sub-optimal response to fluctuations in the natural rate of interest. In particular, we have shown that it results in interest-rate responses that are insufficiently inertial. This raises the question whether it would not be better for a central bank that is expected to act under discretion to pursue an alternative objective, one that includes an interest-rate *smoothing* objective in the sense discussed in the introduction. Note that actual central banks are often described as pursuing interest-rate smoothing as one goal among several, though we argued above that such a goal had no place in the true social welfare function described by (2.8) – (2.9). In fact, such an attitude on the part of central bankers may not be something one should try to correct, if one expects them to pursue their goals in a discretionary fashion.

#### 4.1 Time-Consistent Equilibrium with a Smoothing Objective

We now consider the consequences of delegating the conduct of monetary policy to a central banker that is expected to seek to minimize the expected value of a criterion of the form (2.8), where however (2.9) is replaced by a function of the form

$$L_t^{cb} = \pi_t^2 + \hat{\lambda}_x(x_t - \hat{x}^*)^2 + \hat{\lambda}_r(r_t - \hat{r}^*)^2 + \lambda_\Delta(r_t - r_{t-1})^2. \quad (4.1)$$

Here we allow the weights  $\hat{\lambda}_x, \hat{\lambda}_r$  to differ from the weights  $\lambda_x, \lambda_r$  associated with the true social loss function, and similarly allow the targets  $\hat{x}^*, \hat{r}^*$  to differ from  $x^*$  and  $r^*$ . We also allow for the existence of a term that penalizes interest-rate *changes*, not present in the true social loss function (2.9).

The time-consistent optimizing plan associated with such a loss function can be derived 

---

appointment, or by teaching central bankers that the pursuit of certain objectives rather than others is in the general good.

<sup>62</sup>Kiley (1998) addresses the same issue in the context of a model closely related to the one used here.

using familiar methods, expounded for example in Soderlind (1998). The problem is more complicated than the one considered earlier, because the presence of a term involving the lagged interest rate in the period loss function (4.1) means that even in a Markovian equilibrium, outcomes will depend upon the lagged interest rate. This in turn means that the central bank's expectations at date  $t$  about equilibrium in periods  $t + 1$  and later are not independent of its choice of  $r_t$ . However, in such an equilibrium, the central bank's value function in period  $t$  is given by a function  $V(r_{t-1}; r_t^n)$ , which function is time-invariant. (Here we simplify by assuming that the natural rate of interest is itself a Markovian process, with law of motion (3.15), though we could easily generalize our results to allow for more complicated linear state-space models.)

Standard dynamic programming reasoning implies that the value function must satisfy the Bellman equation<sup>63</sup>

$$V(r_{t-1}; r_t^n) = \min_{(r_t, \pi_t, x_t)} E_t \left\{ \frac{1}{2} [\pi_t^2 + \hat{\lambda}_x (x_t - \hat{x}^*)^2 + \hat{\lambda}_r (r_t - \hat{r}^*)^2 + \lambda_\Delta (r_t - r_{t-1})^2] + \beta V(r_t; r_{t+1}^n) \right\}, \quad (4.2)$$

where the minimization is subject to the constraints

$$\begin{aligned} \pi_t &= \kappa x_t + \beta E_t [\pi(r_t; r_{t+1}^n)], \\ x_t &= E_t [x(r_t; r_{t+1}^n) - \sigma^{-1} (r_t - r_t^n - \pi(r_t; r_{t+1}^n))]. \end{aligned}$$

Here the functions  $\pi(r_t; r_{t+1}^n)$ ,  $x(r_t; r_{t+1}^n)$  describe the equilibrium that the central bank expects to result in period  $t + 1$ , conditional upon the exogenous state  $r_{t+1}^n$ . This represents the consequence of discretionary action at that date and later, that the current central banker regards him or herself as unable to change. Similarly,  $V(r_t; r_{t+1}^n)$  represents the value expected for the central bank's objective as of date  $t + 1$ , in the discretionary equilibrium described by those functions.

Discretionary optimization by the central bank at date  $t$  is then defined by the minimization problem on the right-hand side of (4.2). The solution to this problem, for any given

---

<sup>63</sup>Multiplication of the central bank's loss function by a factor 1/2 here eliminates a factor of 2 from subsequent expressions such as (4.6), and is purely a normalization of the value function.

current state  $(r_{t-1}; r_t^n)$ , defines a set of functions  $r(r_{t-1}; r_t^n)$ ,  $\pi(r_{t-1}; r_t^n)$ ,  $x(r_{t-1}; r_t^n)$ , indicating the optimal values of  $r_t$ ,  $\pi_t$ , and  $x_t$ , and a function  $V(r_{t-1}; r_t^n)$  indicating the minimized value of the right-hand side of the equation. Consistency of the central bank's expectations then requires that the functions  $V$ ,  $\pi$ ,  $x$  used to *define* this minimization problem are identical to the functions obtained as its *solution*.

We shall furthermore restrict attention to solutions of the Bellman equation in which the value function is a *quadratic* function of its arguments, and the solution functions for  $r$ ,  $\pi$ , and  $x$  are each linear functions of their arguments. The solution functions can accordingly be written

$$r(r_{t-1}; r_t^n) = r_0 + r_r r_{t-1} + r_n r_t^n, \quad (4.3)$$

$$\pi(r_{t-1}; r_t^n) = \pi_0 + \pi_r r_{t-1} + \pi_n r_t^n, \quad (4.4)$$

$$x(r_{t-1}; r_t^n) = x_0 + x_r r_{t-1} + x_n r_t^n, \quad (4.5)$$

where  $r_0, r_r, r_n$ , and so on are constant coefficients to be determined by solving a fixed-point problem. The first-order conditions for the optimization problem in (4.2) involve the partial derivative of the value function with respect to the lagged interest rate. This too must be a linear function of its arguments. In fact, differentiation of (4.2) using the envelope theorem implies that when the value function is defined, the partial derivative with respect to its first argument must satisfy

$$V_1(r_{t-1}; r_t^n) = \lambda_\Delta [r_{t-1} - r(r_{t-1}; r_t^n)]. \quad (4.6)$$

Thus linearity of the solution function  $r$  guarantees the linearity of this function as well.

We turn now to the fixed-point problem for the constant coefficients in the solution functions. First of all, substitution of the assumed linear solution functions into the two constraints following (4.2), and using

$$E_t r_{t+1}^n = \rho r_t^n, \quad (4.7)$$

allows us to solve for  $x_t$  and  $\pi_t$  as linear functions of  $r_t$  and  $r_t^n$ . (Let the coefficients on  $r_t$  in the solutions for  $x_t$  and  $\pi_t$  be denoted  $X_r$  and  $\Pi_r$  respectively. These coefficients are

themselves linear combinations of the coefficients  $x_r$  and  $\Pi_r$  introduced in (4.4) – (4.5). Requiring the solution functions defined in (4.3) – (4.5) to satisfy these linear restrictions yields a set of six nonlinear restrictions upon the coefficients  $x_0, x_r, x_n$  and so on.<sup>64</sup>

Substituting these solutions for  $x_t$  and  $\pi_t$  into the right-hand side of (4.2), the expression inside the minimization operator can be written as a function of  $r_t$  and  $r_t^n$ . This expression is quadratic in  $r_t$ , and so it achieves a minimum if and only if it is convex, in which case the optimum is characterized by the first-order condition obtained from differentiation with respect to  $r_t$ . Because the function is quadratic, it is globally convex if and only if the second-order condition is satisfied; thus a solution satisfying the first- and second-order conditions is both necessary and sufficient for optimality.

Substituting (4.6) for the derivative of the value function, the first-order condition may be written as

$$\Pi_r \pi_t + \hat{\lambda}_x X_r(x_t - \hat{x}^*) + \hat{\lambda}_r(r_t - \hat{r}^*) + \lambda_\Delta(r_t - r_{t-1}) + \beta \lambda_\Delta(r_t - E_t r_{t+1}) = 0. \quad (4.8)$$

Substituting into this the above solutions for  $\pi_t$  and  $x_t$  as functions of  $r_t$  and  $r_t^n$ , and the assumed solution (4.3) for  $r_{t+1}$  as a function of  $r_t$  and  $r_t^n$ , we see that the second-order condition may be written as

$$\Omega \equiv \Pi_r^2 + \hat{\lambda}_x X_r^2 + \hat{\lambda}_r + \lambda_\Delta + \beta \lambda_\Delta(1 - r_r) \geq 0. \quad (4.9)$$

Now requiring that the solutions defined in (4.3) – (4.5) always satisfy the linear equation (4.8) gives us another set of three nonlinear restrictions on the constant coefficients of the solution functions. We thus have a set of nine nonlinear equations to solve for the nine coefficients of equations (4.3) – (4.5). A set of coefficients satisfying these equations, and also satisfying the inequality (4.9), represent a linear Markov equilibrium for the central bank objective (4.1).

We shall as usual be interested solely in the case of a *stationary* equilibrium, so that fluctuations in  $r_t, \pi_t$  and  $x_t$  are bounded if the fluctuations in  $r_t^n$  are bounded. (As before, this

---

<sup>64</sup>The restrictions are nonlinear because the coefficients  $X_r$  and so on are themselves functions of the coefficients  $x_r$  and so on.

is the case in which our linear-quadratic approximations are justifiable in terms of a Taylor series approximation to the exact conditions associated with private-sector optimization, in the case of small enough exogenous disturbances.) Given (4.3) – (4.5) and the assumption of stationary fluctuations in  $r_t^n$ , it is clear that  $\pi_t$  and  $x_t$  will be stationary processes as long as  $r_t$  is. It is also obvious that  $r_t$  will be stationary (bounded) if and only if

$$|r_r| < 1. \tag{4.10}$$

Thus we are interested in solutions to the nine nonlinear equations that satisfy both inequalities (4.9) and (4.10).

In the case that  $\hat{\lambda}_x, \hat{\lambda}_r, \lambda_\Delta \geq 0$ , it will be observed that (4.10) implies condition (4.9), so that we need not concern ourselves with the convexity issue in that case. However, non-negativity of these weights in the central-bank objective is not *necessary* for convexity of the central bank’s optimization problem, and it is of some interest to consider delegation to a central banker with a *negative* weight on some term. In particular, we shall see that there are advantages to delegation to a central banker with  $\hat{\lambda}_r < 0$ , while  $\lambda_\Delta > 0$ : the central banker dislikes large interest-rate *changes* of either sign, but actually *prefers* for interest rates to deviate farther from their “target” level  $\hat{r}^*$  (which in such a case is hardly a target!). Such preferences need not result in a violation of convexity, though the negative  $\hat{\lambda}_r$  term itself makes it harder for (4.9) to be satisfied. As long as positive weights are placed on inflation and output variability in the central bank loss function, the effects of current interest-rate decisions on current inflation and output tend to make the central bank’s criterion function convex in  $r_t$ . If in addition  $\lambda_\Delta > 0$ , convexity is increased, both because of the direct effect of the current interest-rate level upon the current interest-rate change, and because of the effect of the current interest-rate decision upon expected future inflation and output. In this case, four out of five terms on the left-hand side of (4.9) are necessarily positive, and the condition will be satisfied as long as these terms together outweigh the negative  $\hat{\lambda}_r$  term.<sup>65</sup>

---

<sup>65</sup>Note that the conditions required for convexity of the bank’s objective at date  $t$ , when it *takes as given* equilibrium outcomes from date  $t + 1$  onward, are much weaker than the conditions that would be required for convexity of its objective if it viewed itself as being able to commit itself at date  $t$  to any state-contingent plan from that date onward that was consistent with the structural equations (2.6) – (2.7).

We turn now to the question of what loss function the central bank should be assigned to minimize, if it is assumed that a time-consistent plan of this sort will be pursued. As above, we begin by considering the steady state associated with such an equilibrium. It is no longer possible, as above, to give a closed-form solution for the steady-state inflation rate without solving for the equilibrium responses to shocks. Nonetheless, it is clear that the parameters  $\hat{x}^*$ ,  $\hat{r}^*$  affect only the steady state, and not the character of the deviations from the steady state resulting from fluctuations in the natural rate. Furthermore, these two parameters of central bank preferences give us more than enough degrees of freedom to bring about the desired steady state, whatever values we may choose for  $\hat{\lambda}_x$ ,  $\hat{\lambda}_r$ , and  $\lambda_\Delta$ . For there is only a one-parameter family of possible steady-state values of  $\pi, x, r$  consistent with the structural equations (2.6) – (2.7), regardless of monetary policy. We thus only need to vary one policy parameter in order to make the equilibrium steady state coincide with the optimal one. For example, we may stipulate that  $\hat{r}^* = r^*$ , the target value in the true social welfare function, and simply adjust  $\hat{x}^*$  to eliminate any bias in the steady-state inflation rate resulting from discretion.<sup>66</sup>

Accordingly, we may determine the desired values of the weights  $\hat{\lambda}_x$ ,  $\hat{\lambda}_r$ , and  $\lambda_\Delta$  simply with regard to achieving desirable equilibrium responses to shocks, taking it as given that the steady state is the optimal one described in (3.26). Then, whatever our choices for the weights, we can choose target values  $\hat{x}^*, \hat{r}^*$  that make the time-consistent steady state coincide with the optimal one.

As a simple example, suppose that  $x^* > 0$ , while  $r^* = 0$ . Then as shown earlier, the optimal steady state involves  $\pi^{opt} = x^{opt} = r^{opt} = 0$ . But this is the steady state consistent with equilibrium conditions (2.6), (2.7) and (4.8) as long as

$$\hat{\lambda}_x X_r \hat{x}^* + \hat{\lambda}_r \hat{r}^* = 0.$$

A sufficient condition for this (though not necessary) would be to choose  $\hat{x}^* = \hat{r}^* = 0$ . Thus, as in the proposal of Blinder (1998, chap. 2), the average inflation bias is eliminated by

---

<sup>66</sup>Because the present specification already has more degrees of freedom than are needed, we have not bothered to consider a target inflation rate different from zero in (4.1).

choosing a central banker whose goal is to stabilize output around the natural rate (*i.e.*, for whom  $\hat{x}^* = 0$ ), even though the true social welfare function is one with  $x^* > 0$ .<sup>67</sup>

We turn now to the consequences of the weights in the central-bank loss function (4.1) for the equilibrium responses to shocks. As conjectured above, we find that setting  $\lambda_\Delta > 0$  results in inertial interest-rate responses to fluctuations in the natural rate. If we take the partial derivative of the left-hand side of (4.8) with respect to  $r_{t-1}$ , using the solution functions to express all terms as functions of  $r_{t-1}$  and  $r_t^n$ , we obtain a coefficient equal to  $\Omega r_r - \lambda_\Delta$ , where  $\Omega$  is defined in (4.9). The first-order condition (4.8) thus implies that

$$\Omega r_r = \lambda_\Delta$$

in any solution. Then if  $\lambda_\Delta > 0$ , both  $\Omega$  and  $r_r$  must be non-zero, and of the same sign. The second-order condition (4.9) then implies that in any equilibrium, both quantities must be positive. Combining this result with (4.10), we conclude that in any stationary equilibrium,

$$0 < r_r < 1. \tag{4.11}$$

Given this, the law of motion (4.3) for the nominal interest rate implies partial adjustment toward a time-varying target that is a linear function of the current natural rate of interest. Since this is at least a rough characterization of a way in which the optimal responses to shocks differ from the time-consistent responses when the central bank seeks to minimize true social losses, it is plausible that delegating monetary policy to a central banker who believes it is better to reduce the variability of interest-rate changes can improve social welfare.

One may wonder whether it is possible to choose the weights in the central bank's loss function so as to completely eliminate the distortions associated with discretion, and achieve the same responses as under an optimal commitment. It should be immediately apparent that it is not in general possible to achieve this outcome exactly. For we have shown above

---

<sup>67</sup>Of course, as in Rogoff (1985), one might alternatively keep  $\hat{x}^*$  equal to the target value  $x^*$  in the true social welfare function, and instead seek to mitigate the inflation bias by adjusting the value of  $\hat{\lambda}_x$ . The inflation bias could be completely eliminated by setting  $\hat{\lambda}_x = 0$ . However, the value of  $\hat{\lambda}_x$  also affects the equilibrium responses to shocks, and such distortion of the weight put on output gap stabilization may have undesirable consequences for that reason. Indeed, Rogoff argues that it is not desirable to adjust  $\hat{\lambda}_x$  to the point that the inflation bias is fully eliminated, exactly because of this tension.

that the optimal interest-rate dynamics have a representation of the form (3.32), where  $Q(L)$  is of second order and  $R(L)$  is of first order; thus they generally do not take a form as simple as (4.3). Nonetheless, we can show that exact implementation of the optimal plan is possible at least in a limiting case. And we can also show that it is possible to achieve a pattern of responses nearly as good as the optimal plan, in the “calibrated” numerical example taken up earlier. These points are taken up in succession in the next two subsections.

## 4.2 Optimal Delegation in a Limiting Case

Here we consider again the limiting case with  $\kappa = 0$  taken up in section 3.5. We have shown there that in this special case, the optimal interest-rate and output dynamics do take the form given by (4.3) and (4.5), where the coefficients (neglecting the constant term) are given by

$$r_r = \mu_1, \quad r_n = (1 - \mu_1) \left( \frac{\mu_2 - 1}{\mu_2 - \rho} \right), \quad (4.12)$$

$$x_r = -(\beta^{-1} - \mu_1) \sigma \frac{\lambda_r}{\lambda_x}, \quad x_n = \frac{1}{\sigma(\mu_2 - \rho)}. \quad (4.13)$$

Here  $\mu_1, \mu_2$  refer to the two roots of (3.34) discussed earlier. The question that we wish to ask, then, is whether it is possible to choose the weights in (4.1) so that the optimal values (4.12) – (4.13) solve the equilibrium conditions just derived. Since the optimal values necessarily satisfy the conditions required for consistency with the structural equation (2.7), it suffices that they also be consistent with conditions (4.8) and (4.9) for time-consistent optimizing behavior on the part of the central bank.

In this limiting case, the conditions required for consistency of (4.3) and (4.5) with (4.8) simplify to

$$(x_r - \sigma^{-1}) \hat{\lambda}_x x_r + \hat{\lambda}_r r_r + (1 - \beta r_r) \lambda_\Delta (r_r - 1) = 0,$$

$$(x_r - \sigma^{-1}) \hat{\lambda}_x x_n + \hat{\lambda}_r r_n + (1 + \beta(1 - r_r - \rho)) \lambda_\Delta r_n = 0.$$

These two equations depend only upon the ratios of the weights in the policy objective,  $\hat{\lambda}_r/\hat{\lambda}_x$  and  $\lambda_\Delta/\hat{\lambda}_x$ , rather than upon the absolute size of the three weights. (This is because inflation

variations are negligible under any policy regime, so the relative weight on inflation variability no longer matters.) Hence we may, without loss of generality, suppose that  $\hat{\lambda}_x = \lambda_x$ , the weight in the true social objective function. Using this simplification, and substituting the optimal values (4.12)–(4.13), the above two linear equations can be solved for the unique values of  $\hat{\lambda}_r$  and  $\lambda_\Delta$  consistent with the optimal equilibrium responses. These are given by

$$\lambda_\Delta = \lambda_r \frac{\lambda_r \sigma^2 (\beta^{-1} - \mu_1) + \lambda_x}{(1 - \beta \rho \mu_1) \beta \lambda_x} > 0, \quad (4.14)$$

$$\hat{\lambda}_r = -(1 - \beta \rho)(1 - \beta \mu_1) \lambda_\Delta < 0. \quad (4.15)$$

While one finds that the kind of partial-adjustment interest-rate dynamics associated with the optimal plan do require  $\lambda_\Delta > 0$ , as conjectured, one finds that they cannot be exactly matched through delegation to a central banker with discretion unless in addition  $\hat{\lambda}_r < 0$ . This is another difference between the best objective for the central banker and the true social objective function, since in the latter,  $\lambda_r > 0$ . As noted earlier, a negative value for  $\hat{\lambda}_r$  does not necessarily imply violation of the convexity condition (4.9) needed for central-bank optimization. In fact, we have shown above that the convexity condition holds in the case of any solution to the first-order condition with  $\lambda_\Delta > 0$  and  $r_r > 0$ . As (4.12) implies that  $r_r > 0$ , and (4.14) implies that  $\lambda_\Delta > 0$ , the above assumed central-bank objective does result in a convex optimization problem for the central bank. Thus the optimal pattern of responses to shocks can in this case be supported as an equilibrium outcome under discretion, as long as the central bank is charged with pursuit of an objective that involves interest-rate smoothing.

### 4.3 A Numerical Example

When  $\kappa > 0$ , no such simple analytical result is available. It might in this more general case still be possible to support the optimal plan as an equilibrium outcome under discretion, if a central-bank objective more complex than (4.1) were considered. We do not pursue this here, but instead note that even with an objective in the simpler class (4.1), it is possible to achieve quite a good approximation to the optimal pattern of responses to shocks, in the

case of plausible parameter values. We demonstrate this through numerical analysis of the model, assuming the same “calibrated” parameter values as earlier.

Specifically, we assume that the parameters of the structural equations and the shock process are as specified in Table 1. But we now assume a central-bank loss function of the form (4.1), and consider how the time-consistent optimizing plan varies with the assumed weights in that loss function. To begin, we shall assume that  $\hat{\lambda}_x = \lambda_x = .048$  (the value in Table 1), and consider only the consequences of variation in  $\hat{\lambda}_r$  and  $\lambda_\Delta$ .

We first note that the nonlinear equations referred to above do not always have a unique solution for the coefficients  $r_r, r_n$ , and so on. It can be shown that given a value for  $r_r$  consistent with these equations, a unique solution can be obtained, generically, for the other coefficients. However,  $r_r$  solves a quintic equation, which equation may have as many as five real roots. For example, Figure 7 plots the solutions to this equation, as a function of  $\hat{\lambda}_r$ , in the case that  $\lambda_\Delta = 0$ . One observes that there is a unique real root,  $r_r = 0$ , in the case of any  $\hat{\lambda}_r > 0$ .<sup>68</sup> However, for  $\hat{\lambda}_r < 0$ , there are multiple solutions, and given the results of the previous sub-section, we are interested in considering loss functions of this kind. In the figure, solutions that also satisfy conditions (4.9) and (4.10), and so correspond to stationary equilibria, are indicated by solid lines, while additional branches of solutions that do not correspond to stationary equilibria are indicated by dashed lines.<sup>69</sup> We observe that while there exist multiple solutions to the nonlinear equations for all  $\hat{\lambda}_r < 0$ , there is still a unique stationary equilibrium involving optimization under discretion for all  $\hat{\lambda}_r > -1$ .<sup>70</sup>

---

<sup>68</sup>The time-consistent solution characterized in section 3.2 refers to a case of this sort. In such a case, the unique solution involves  $r_r = \pi_r = x_r = 0$ , so that lagged interest rates have no effect. This solution obviously satisfies (4.9) and (4.10) as well, and so represents the unique stationary Markov equilibrium.

<sup>69</sup>Technically, because here  $\lambda_\Delta = 0$ , the second-order condition is (weakly) satisfied even by solutions in which  $r_r < 0$ . But our real interest is not in the case  $\lambda_\Delta = 0$ , but rather in the set of solutions that exist for small positive values of  $\lambda_\Delta$ . The solutions shown in Figure 7 with  $r_r < 0$  also correspond to solutions with  $r_r < 0$  in the case of small positive  $\lambda_\Delta$ , and under that perturbation these solutions cease to satisfy the second-order condition. Hence we show these branches of solutions with dashed lines in Figure 7. The correct statement would be that for any *small enough* value  $\lambda_\Delta > 0$ , there exists a unique stationary equilibrium for all  $\hat{\lambda}_r > -1$ . This identifies the boundary of the white region in Figures 8-10, near the horizontal axis.

<sup>70</sup>It is interesting to note that for values of  $\hat{\lambda}_r$  below a critical value, approximately -0.02, the unique stationary equilibrium no longer corresponds to the “minimum state variable solution”, *i.e.*, the solution in which lagged interest rates are irrelevant.

Only for even larger negative values do we actually have multiple time-consistent equilibria. The same turns out to be true for  $\lambda_\Delta > 0$  as well, at least in the case of the moderate values of  $\lambda_\Delta$  that we shall consider here. (For very high values of  $\lambda_\Delta > 0$ , not shown in the figures below, there exist multiple equilibria even for higher values of  $\hat{\lambda}_r$ .) Note that our time-consistent outcome is essentially a Nash equilibrium in a game played by successive central bankers, rather than the solution to an optimization problem. Thus there is nothing paradoxical about the possible existence of multiple solutions.

We turn now to a consideration of how the properties of the stationary time-consistent equilibrium vary with the parameters  $\hat{\lambda}_r$  and  $\lambda_\Delta$ . In each of Figures 8-10, the white region indicates the set of loss function weights for which there is a unique stationary equilibrium of the linear form characterized above. In this region, the contour lines plot properties of this equilibrium. The grey region indicates weights for which there are multiple stationary equilibria. Here we plot the values associated with the best of these equilibria, the one with the lowest value of  $E[W]$ . As it turns out, the best equilibrium that is attainable corresponds to weights in the white region, so that we do not have to face the question of whether one should choose weights that are consistent with *one* good equilibrium but also with other bad ones.

Figure 8 shows how the inertia coefficient  $r_r$  in representation (4.3) of the equilibrium interest-rate dynamics varies with the loss function weights. As one might expect, the equilibrium inertia coefficient increases as  $\lambda_\Delta$  is increased, for any given value of  $\hat{\lambda}_r$ . At the same time, for any given value of  $\lambda_\Delta > 0$ , the inertia coefficient also increases if  $\hat{\lambda}_r$  is reduced. This continues to be true as  $\hat{\lambda}_r$  is made negative. Figure 9 shows the corresponding equilibrium values of the coefficient  $r_n$ , describing the immediate interest-rate response to an increase in the natural rate of interest. Increasing either  $\hat{\lambda}_r$  or  $\lambda_\Delta$  lowers this response coefficient, at least in the region where both are positive, though the response is positive.

Figure 10 presents the corresponding contour plot for the true social loss measure  $E[W]$ . Four sets of policy weights are marked on this figure (as on the others). The X indicates the weights in the true social loss function; but charging a discretionary central bank to minimize

this objective does not lead to the best equilibrium, even under this criterion. (These weights, and properties of the resulting time-consistent equilibrium, are described in the first line of Table 3.) The large black dot instead indicates the weights that lead to the best outcome, when one still restricts attention to central bank loss functions with no smoothing objective ( $\lambda_\Delta = 0$ ). This corresponds to a weight  $\hat{\lambda}_r$  that makes the time-consistent equilibrium implement the optimal *non-inertial* plan, characterized earlier. (Compare the second line of Table 3 with the fifth line of Table 2.) It involves a value  $\hat{\lambda}_r < \lambda_r$ , so that interest rates respond more vigorously to variations in the natural rate of interest than occurs under discretion when the central bank seeks to minimize the true social loss function.

The circled star (or wheel) instead indicates the minimum achievable value of  $E[W]$ , among time-consistent equilibria of this kind. These weights therefore solve the optimal delegation problem, if we restrict ourselves to central-bank objectives of the form (4.1). As in the limiting case solved explicitly above, the optimal weights involve  $\lambda_\Delta > 0, \hat{\lambda}_r < 0$ . (The optimal weights and the properties of the resulting equilibrium are described on the fourth line of Table 3.) Finally, the star without a circle indicates the optimal equilibrium that can be achieved subject to the constraint that  $\hat{\lambda}_r \geq 0$ . (One might restrict attention to these cases, if one does not think it would be easy to explain to a central bank that it will be rewarded for creating interest-rate variability, while being punished for failing to smooth interest-rate changes.) This point corresponds to a point of tangency between an isoquant of  $E[W]$  and the vertical axis at  $\hat{\lambda}_r = 0$  (though neither curve is drawn in the figure). In this case, it is still desirable to direct the central bank to penalize large interest-rate changes, though the optimal  $\lambda_\Delta$  is smaller than if it were possible to choose  $\hat{\lambda}_r < 0$ . (These constrained-optimal weights are shown on the third line of Table 3.)

In this exercise, we have assumed that the relative weight on output gap variability,  $\hat{\lambda}_x$ , equals the weight in the true social loss function,  $\lambda_x$ , given in Table 1. In fact, consideration of values  $\hat{\lambda}_x \neq \lambda_x$  allows us to do no better, either in the case of loss functions with no smoothing objective, or in the case of the fully unconstrained family. For the policy on the second line of Table 3 already implements the optimal non-inertial plan, and the policy on the

fourth line already implements the optimal plan of the form given by equations (4.3) – (4.5). This is not because  $\hat{\lambda}_x = \lambda_x$  is a uniquely optimal value in either case, but rather because we can find weights that support the optimal plan for an arbitrary value of  $\hat{\lambda}_x$ , so that the constraint that  $\hat{\lambda}_x = \lambda_x$  has no cost. The structural equations (2.6) and (2.7) place four restrictions upon the coefficients  $(r_r, r_n, \pi_r, \pi_n, x_r, x_n)$ , so that there is only a two-parameter family of possible responses to shocks that are consistent with the structural equations. To pick out the optimal member of this family, it suffices that one be able to freely vary two parameters of the central bank loss function, such as  $\hat{\lambda}_r$  and  $\lambda_\Delta$ . For example, it would also be possible to impose the constraint that  $\hat{\lambda}_x = 0$ , so that there is no output-gap term in the central bank loss function at all. The optimal weights in this case are given on the fifth line of Table 3. Note that again  $\hat{\lambda}_r < 0$ ,  $\lambda_\Delta > 0$ .

The only case in which it is of some benefit to relax the requirement that the relative weight on the inflation and output-gap terms in (4.1) are the same as in the true social loss function is when we impose the constraint that all weights be non-negative. As noted above, in this case the constraint that  $\hat{\lambda}_r \geq 0$  binds. But if we must set  $\hat{\lambda}_r = 0$ , the additional degree of freedom allowed by varying  $\hat{\lambda}_x$  as well as  $\lambda_\Delta$  does allow some improvement of the time-consistent equilibrium, in general. In fact, for the numerical parameter values used above, the optimal  $\hat{\lambda}_x$  is infinite; that is, the relative weight on the inflation term is best set to zero. To analyze this case, it is thus convenient to instead write the central bank loss function as

$$L_t^{cb} = (x_t - \hat{x}^*)^2 + \tilde{\lambda}_\pi \pi_t^2 + \tilde{\lambda}_r (r_t - \hat{r}^*)^2 + \tilde{\lambda}_\Delta (r_t - r_{t-1})^2. \quad (4.16)$$

In terms of this alternative notation, the loss function described on the fourth line of Table 3 can instead be described as on the sixth line of the Table. We now consider the optimal central bank of objective of the form (4.16), when we impose the constraint that  $\tilde{\lambda}_r \geq 0$ .

Figure 11 plots the value of  $E[W]$  in the time-consistent equilibrium, as a function of the policy weights  $\tilde{\lambda}_\pi$  and  $\tilde{\lambda}_\Delta$ , in the case that  $\tilde{\lambda}_r = 0$ . As in the earlier figures, the X marks the weights in the true social loss function (corresponding to the first line in Table 3). The star without a circle again indicates the best attainable policy in this plane, under the constraint

that  $\tilde{\lambda}_\pi = 1/\lambda_x$ , the weight in the true social loss function. The star is at a point of tangency between an  $E[W]$  isoquant and a vertical line at  $\tilde{\lambda}_\pi = 1/\lambda_x$ , and corresponds to the policy described on the third line of Table 3. However, we observe that it is possible to reduce  $E[W]$  by lowering  $\tilde{\lambda}_\pi$ , and the minimum value consistent with non-negative policy weights (indicated by the wheel in the figure, and described on the seventh line of Table 3) involves  $\tilde{\lambda}_\pi = 0$ , so that a positive weight is assigned to output-gap stabilization, and *none* to inflation stabilization.<sup>71</sup> Note that this constrained-optimal central bank loss function again involves  $\tilde{\lambda}_\Delta > 0$ , so that it is desirable for the central bank to seek to smooth interest-rate changes.

How good an equilibrium can be achieved through optimal delegation of this kind? One will recall from Table 2 that for these parameter values (the case  $\rho = .35$  in that table), the optimal policy achieves an expected loss of  $E[W] = 1.097$ , while the expected loss is more than twice as large ( $E[W] = 2.547$ ) under the time-consistent optimizing plan (assuming a central bank that seeks to minimize the true social loss measure). However, the minimum value of  $E[W]$  shown in Figure 10 is also equal to 1.097, to three significant digits. Thus more than 99.9 percent of the reduction in expected loss that is possible in principle, through an optimal commitment, can be achieved through an appropriate choice of objective for a discretionary central bank.<sup>72</sup> The exact optimal pattern of responses could presumably be supported as a time-consistent equilibrium if we were to consider more complex central bank loss functions.

## 5 Optimal Interest-Rate Feedback Rules

We turn now to an alternative approach to implementation of the optimal pattern of responses to shocks. Here we ask what kind of policy rule would achieve that end, assuming

---

<sup>71</sup>Further insight into why it is optimal to set  $\tilde{\lambda}_\pi = 0$  in this case may be provided by Figure 12, which plots  $E[W]$  as a function of  $\tilde{\lambda}_r$  and  $\tilde{\lambda}_\pi$ , fixing  $\tilde{\lambda}_\Delta$  at the value shown on the sixth line of Table 3, *i.e.*, the unconstrained optimal weight. The wheel again indicates the unconstrained optimal central bank loss function. Under the constraint that  $\tilde{\lambda}_r \geq 0$ , however, the best attainable point in this plane is at the star, corresponding to  $\tilde{\lambda}_\pi = \tilde{\lambda}_r = 0$ .

<sup>72</sup>The equilibrium achieved in this way is also very similar in other respects, such as the other statistics for the optimal plan reported in Table 2.

that the central bank commits itself to systematically conduct policy in this particular way, and that the commitment is understood by and credible to the private sector. We shall assume that the kind of characterization of optimal policy that is sought is a *feedback rule* expressing the short-run nominal interest rate (the central bank's instrument) as a function of current and lagged values of various variables that may be observed by the central bank, and lagged values of the interest rate itself, as in the econometric studies mentioned in the introduction. The question in which we are especially interested is the nature of the dependence upon lagged interest rates in an optimal rule.

It is important to recognize that the question of optimal interest-rate inertia in this sense is a distinct one from the question of the optimal response of interest rates to shocks taken up in section 3. This is because a mere description of how one would like for the interest rate to vary as a function of the history of shocks does not in itself uniquely identify a feedback rule for the central bank that achieves this outcome. We have obtained one description of interest-rate dynamics under an optimal plan, given by equation (3.32). But this need not be the central bank's *policy rule*, since many other relations between interest rates and other state variables are equally correct descriptions of a relation that should hold under the optimal plan. If the variables in question are all part of the central bank's information set, then there is no reason not to suppose that another of these relations might equally well represent an appropriate policy rule.

One might think that the relation (3.32) is a uniquely appropriate way of representing optimal policy, since it is the unique relation that holds in equilibrium that involves *only* the interest rate itself and exogenous states. However, there are two important reasons why other relations satisfied by the nominal interest rate under the optimal plan may be of greater usefulness as proposed policy rules. One is that other rules that are equally consistent with the optimal equilibrium may require *less information* for implementation by the central bank. Relation (3.32), if proposed as a policy rule, would require the central bank to determine at each point in time the current value of the natural rate of interest. But as we shall see, it is possible for the central bank to adopt a feedback rule that involves no explicit dependence

upon the natural rate, and so does not even require the central bank to know the value of that variable, and yet results in the desired rational expectations equilibrium variation of inflation and interest rates with disturbances to the natural rate. Under such a regime, the central bank relies upon the *private sector's* awareness of the current underlying state variables to bring about the desired responses of endogenous variables to these states.

The second reason is that rules that are equally *consistent* with the optimal equilibrium may not serve equally to *uniquely determine* the optimal plan as the equilibrium outcome. This is because rational expectations equilibrium may be *indeterminate* under some interest-rate feedback rules that include the optimal plan as one of the many possible equilibria. In fact, the relation (3.32), if adopted as a *rule* for fixing the central bank's interest-rate target each period, would have this unfortunate feature. But as we shall see, other relatively simple feedback rules exist which result in a *determinate* equilibrium (in the sense that there is at any rate a unique *stationary* equilibrium) which achieves the optimal plan. Such rules are accordingly of greater interest as candidate monetary policy rules.

## 5.1 The Problem of Determinacy of Equilibrium

A fundamental issue in the evaluation of alternative policy rules in a rational expectations equilibrium framework, when one's structural model possesses forward-looking elements, is the question of whether a proposed policy rule is associated with a determinate equilibrium or not. Policy rules may easily be associated with very large sets of rational expectations equilibria, and this problem of *indeterminacy* is an important reason for excluding certain categories of rules from consideration.<sup>73</sup> Sargent and Wallace (1975) famously argued for money-supply rules as opposed to interest-rate rules for monetary policy on the ground that interest-rate rules resulted in price-level indeterminacy. In fact, as McCallum (1981) showed, this is not a problem of interest-rate rules *as such*, even in the context of the particular rational expectations IS-LM model used by Sargent and Wallace. Rather, their

---

<sup>73</sup>See, e.g., Bernanke and Woodford (1997), Rotemberg and Woodford (1998), Clarida *et al.* (1998a), Schmitt-Grohé and Uribe (1999), Christiano and Gust (1999), and Woodford (1999a) for recent examples of analyses of monetary policy rules that emphasize this issue.

result applies only to interest-rate rules that specify each period's nominal interest rate as a function solely of exogenous states. However, this particular result applies to our structural model as well, and it explains why the law of motion (3.32) for the nominal interest rate is not a suitable interest-rate *rule* through which to implement optimal policy.

Consider any policy rule in which the central bank's instrument  $r_t$  evolves exogenously. We can represent such a rule by an equation of the form

$$\hat{r}_t = \xi' s_t, \quad (5.1)$$

where  $s_t$  is the vector of exogenous states that evolves according to (3.12), and  $\xi$  is a vector of coefficients. (Here we suppose that the vector  $s_t$  may be extended to include states that are not needed to forecast the evolution of the natural rate of interest, but that matter for the central bank's setting of its instrument. In particular, we may now allow the vector  $s_t$  to include lagged values of the interest rate itself, so that rules of the form (3.32) may be given a representation of the form (5.1).) Then using (5.1) to substitute for  $r_t$  in (2.7), our structural equations take the form

$$E_t z_{t+1} = \hat{A} z_t + \hat{a} s_t, \quad (5.2)$$

where

$$\hat{A} \equiv \begin{bmatrix} \beta^{-1} & -\beta^{-1}\kappa \\ -\beta^{-1}\sigma^{-1} & 1 + \kappa\beta^{-1}\sigma^{-1} \end{bmatrix}, \quad \hat{a} \equiv \begin{bmatrix} 0 \\ \sigma^{-1}(\xi' - k') \end{bmatrix}.$$

Given the signs assumed for  $\beta$ ,  $\sigma$ , and  $\kappa$ , it is apparent that the matrix  $\hat{A}$  has two real eigenvalues, satisfying

$$0 < \lambda_1 < 1 < \lambda_2.$$

The fact that one eigenvalue has modulus less than one, even though neither element of  $z_t$  is a predetermined variable, implies that rational expectations equilibrium is indeterminate in this case. Specifically, for  $\{z_t\}$  any bounded stochastic process satisfying (5.2), another bounded solution is given by

$$z'_t = z_t + v\psi_t,$$

where  $v$  is the right eigenvector of  $\hat{A}$  associated with the eigenvalue  $\lambda_1$ , and the stochastic process  $\psi_t$  evolves according to

$$\psi_{t+1} = \lambda_1 \psi_t + \nu_{t+1},$$

where  $\nu_{t+1}$  is any bounded random variable with mean zero that is unforecastable at date  $t$ , and the initial condition  $\psi_0$  may be chosen arbitrarily. Note that the disturbances  $\nu_t$  may be of any amplitude,<sup>74</sup> and may have any correlation with the “fundamental” disturbances  $\epsilon_t$ . Thus the system (5.2) admits a large multiplicity of bounded (or stationary) solutions, including both solutions implying different equilibrium responses to shocks to “fundamentals” and solutions involving responses to “sunspot” states as well. Furthermore, these solutions include equilibria involving arbitrarily large fluctuations in the endogenous variables  $\pi_t$  and  $x_t$ . Hence even if the rule is chosen (as in the case of (3.32)) to be consistent with the optimal equilibrium, the set of rational expectations equilibria consistent with such a rule includes equilibria that are very bad, from the point of view of the welfare criterion (2.8) – (2.9). Thus such a rule does not represent a desirable way of implementing optimal policy.

## 5.2 An Optimal Simple Rule

The problem of indeterminacy of rational expectations equilibrium can be resolved through sufficiently strong feedback from endogenous variables, such as the inflation rate, the output gap, or both.<sup>75</sup> (Note that while McCallum (1981) stresses the role of feedback from a “nominal anchor” in the resolution of the problem of price-level indeterminacy, no feedback from any nominal variable is actually necessary, given the link between real and nominal variables implied by our aggregate supply relation.) However, once we consider the possibility of feedback from additional variables, there ceases to be a *unique* set of coefficients for the policy rule that is consistent with the optimal equilibrium. Even when we add the additional

---

<sup>74</sup>This is subject, of course, to the caveat that our log-linear approximation to the structural equations may break down in the case of fluctuations that are too large.

<sup>75</sup>Because our system of structural equations (2.6) – (2.7) involves only these two endogenous variables, in addition to the interest rate itself, we consider only rules involving feedback from these variables here. Consideration of the effects of feedback from other variables would require that we adjoin to our model additional equations for determination of those variables.

requirement that the policy rule support the optimal plan as the *unique* bounded rational expectations equilibrium, the set of policies that satisfy this requirement will be large, since the additional requirement of determinacy adds only a set of *inequality* constraints that the coefficients must satisfy.

This allows us to consider additional desiderata in choosing a policy rule. An additional criterion of considerable practical interest is minimization of the information required in order for the central bank to implement the rule in question. From this point of view, a rule that involves feedback only from the inflation rate (*i.e.*, involving no dependence upon either exogenous states or the output gap) is especially desirable, since it can be implemented without the central bank having to possess good estimates of the current values of either the natural rate of interest or the natural rate of output. The uniquely minimal information requirements of such a rule favor both accuracy of implementation and *transparency*, so that the chances are greatest that the private sector will be able to correctly anticipate future monetary policy. A rule that is purely *backward-looking*, responding only to inflation that has already occurred rather than to forecasts of future inflation (whether those of the central bank itself or of outside parties) is also advantageous on these same grounds.

As an illustration of these principles, let us consider policy rules within the simple family

$$\hat{r}_t = \theta \hat{r}_{t-1} + \phi_\pi \hat{\pi}_t + \phi_x \hat{x}_t + \phi_n r_t^n. \quad (5.3)$$

Note that this family includes the sort of generalized “Taylor rule” with partial-adjustment dynamics described by equations (1.1)–(1.2), though we shall not necessarily assume that  $\theta < 1$ , as would be required by the partial-adjustment interpretation. We also allow for the possibility of direct feedback from an observation of the current natural rate of interest, if available. For simplicity, we also again assume a first-order autoregressive process (3.15) for the natural rate of interest. Our complete set of structural equations then consists of (2.6) – (2.7), (3.15), and (5.3).

This system (setting aside the equation for the exogenous dynamics of the natural rate)

can be written in matrix form as

$$\begin{bmatrix} E_t z_{t+1} \\ \hat{r}_t \end{bmatrix} = \begin{bmatrix} \hat{A} & -\theta a \\ \phi' & \theta \end{bmatrix} \begin{bmatrix} z_t \\ \hat{r}_{t-1} \end{bmatrix} + \begin{bmatrix} \hat{a} \\ \phi_n \end{bmatrix} r_t^n, \quad (5.4)$$

where  $\phi' \equiv [\phi_\pi \phi_x]$  is the row vector of coefficients indicating the policy feedback from the endogenous variables  $z_t$ ,  $\hat{A}$  is a  $2 \times 2$  matrix the elements of which depend upon the coefficients  $\phi$ ,<sup>76</sup>  $a$  is the vector that also appears in (3.11), and  $\hat{a} \equiv (1 - \phi_n)a$ . Because there is one predetermined endogenous variable ( $\hat{r}_{t-1}$ ), the system (5.4) has a unique bounded solution if and only if exactly two eigenvalues of the  $3 \times 3$  matrix lie outside the unit circle. The conditions for this to obtain imply a set of inequalities that the coefficients  $\phi_\pi$ ,  $\phi_x$  and  $\theta$  must satisfy (that also involve the structural parameter  $\beta, \sigma$ , and  $\kappa$ . We shall not here develop these conditions, but we note that for any value  $|\theta| < 1$ , the required conditions are violated for all  $\phi$  in a neighborhood of zero, as we should expect from our above analysis of an exogenous interest-rate target.<sup>77</sup> We also note that a sufficient condition for determinacy is that  $\phi_\pi, \phi_x, \theta \geq 0$ , and that

$$\phi_\pi + \theta > 1. \quad (5.5)$$

This last condition expresses the ‘‘Taylor principle’’ that if inflation persists at a level above its long-run target level for a sufficient period of time, the interest rate is eventually raised by *more* than the amount by which inflation exceeds its target level.

In the case that equilibrium is determinate, we shall again be able to solve for the non-predetermined endogenous variables  $z_t$  as a linear function of the predetermined and exogenous state variables  $r_{t-1}$  and  $r_t^n$ . (Here we use the assumption that  $r_t^n$  is Markovian, so that the only relevant exogenous state is  $r_t^n$ .) It then follows from (5.3) that  $\hat{r}_t$  will also be a linear function of the same two variables. Thus the equilibrium dynamics are necessarily of the form (4.3) – (4.5). It follows that at best, a rule within the family (5.3) may be

---

<sup>76</sup>Note that this matrix reduces to the matrix  $\hat{A}$  in (5.2) in the case that the coefficients of the policy rule are  $\phi = 0$ , and to the matrix  $A$  in (3.11) in the case that the coefficients  $\phi$  are those of (3.10).

<sup>77</sup>Note that when  $\phi = 0$ , the eigenvalues of the matrix in (5.4) reduce to the two eigenvalues of  $\hat{A}$ , characterized in section 5.1, and  $\theta$ . As one eigenvalue of  $\hat{A}$  satisfies  $0 < \lambda_1 < 1$ , there are too many eigenvalues inside the unit circle for determinacy. By continuity, the same is true for all coefficients  $\phi$  in a neighborhood of zero.

chosen so as to achieve the optimal pattern of responses to shocks among those that can be described by partial-adjustment dynamics of this form. But we have seen earlier, in sections 3.5 and 3.6, that partial-adjustment dynamics of this kind can be found that approximate the optimal impulse responses reasonably well, as long as  $\kappa$  is small. Furthermore, we have already identified the optimal responses from among this family, as they are the ones associated with the time-consistent equilibrium when the policy weights in the central bank loss function (4.1) are chosen optimally (that is, the responses in the equilibria described in lines 4 – 6 of Table 3). Thus we turn to the question of whether a feedback rule of the form (5.3) can be found that achieves this near-optimal pattern of responses as a determinate equilibrium.

The near-optimal pattern of responses described by the coefficients  $(r_r, r_n, \pi_r, \pi_n, x_r, x_n)$  is *consistent* with the policy rule (5.3) if and only if the coefficients of the rule satisfy two linear restrictions:

$$r_r = \theta + \phi_\pi \pi_r + \phi_x x_r, \quad (5.6)$$

$$r_n = \phi_\pi \pi_n + \phi_x x_n + \phi_n. \quad (5.7)$$

Thus there will be a two-parameter family of such rules, each of which is equally consistent with the desired equilibrium.

We could, of course, obtain a unique selection from this two-parameter family by stipulating that  $\phi_\pi = \phi_x$ , so that there is no feedback from endogenous variables. But in this case, we have seen that the conditions for determinacy of equilibrium will not be met. On the other hand, imposing the requirement of determinacy does not in itself result in a unique selection, because this only imposes additional inequality constraints. Thus there remain, in general, a set of equally satisfactory rules corresponding to an open subset of the two-dimensional linear space of coefficients that satisfy (5.6)–(5.7).

A further possible criterion for a desirable policy rule, as mentioned above, is that it require less information for the central bank to implement it. In the case of the family of rules (5.3), the central bank can implement the rule without having to observe the current exogenous disturbances  $r_t^n$  and  $y_t^n$  if and only if  $\phi_x = \phi_n = 0$ . (We assume here that the

bank can observe inflation and output, but that it cannot observe the current value of the *output gap* without also being able to observe the current natural rate of output.) This criterion thus provides two more linear restrictions upon the coefficients of the policy rule, sufficient to identify a *unique* rule that is both consistent with the desired equilibrium and implementable without any need to observe the exogenous shocks directly. The unique rule of this kind is a rule of the form

$$\hat{r}_t = \theta \hat{r}_{t-1} + \phi_\pi \hat{\pi}_t, \quad (5.8)$$

with coefficients

$$\theta = r_r - r_n \frac{\pi_r}{\pi_n}, \quad \phi_\pi = \frac{r_n}{\pi_n}. \quad (5.9)$$

In the case of parameter values like those investigated in section 3.5, the optimal impulse responses to fluctuations in the natural rate of interest are well approximated by partial-adjustment dynamics with coefficients

$$\begin{aligned} 0 < r_n < 1, & & 0 < r_r < 1, \\ \pi_n > 0, & & \pi_r < 0. \end{aligned}$$

This is because the nominal interest rate adjusts gradually toward a moving target that is an increasing function of (but moves less than proportionally with) the natural rate of interest; while the inflation rate increases at the time of a positive innovation in the natural rate, but subsequently undershoots its long-run level. In such a case, the coefficients (5.9) satisfy

$$\phi_\pi > 0, \quad \theta > r_r.$$

Such a rule results in a determinate rational expectations equilibrium, involving the desired near-optimal pattern of responses to shocks, if the coefficients  $\phi_\pi$  and  $\theta$  are not just positive, but large enough to satisfy (5.5). In fact, for empirically realistic parameter values, each coefficient is likely to be greater than one on its own, so that the determinacy condition is easily satisfied.

Thus we again conclude that intrinsic inertia in interest rates is justified, here measured by the positive coefficient  $\theta$  in the feedback rule (5.8). In fact, we find not only that a

positive  $\theta$  is called for, but one greater than  $r_r$ . Thus the degree of intrinsic inertia that should be incorporated into a feedback rule of this kind is *even greater* than the degree of inertia previously determined to be optimal in the dynamic response of the interest rate to exogenous disturbances. The reason is that under a policy of the kind described by (5.8), the central bank does not respond directly to exogenous disturbances, but instead reacts to them only indirectly, as a result of their effects upon inflation. In this case, there is an additional reason for the interest rate at any time to be an increasing function of the previous period's interest rate, apart from the desire for inertia in the response to variations in the natural rate. This is that the disturbance to the natural rate of interest that can be inferred from any given level of inflation *is higher if nominal interest rates have been higher in the recent past*. This is because, in equilibrium, higher nominal interest rates are followed by lower inflation; so the degree of exogenous inflationary pressure that can be inferred from any given level of current inflation is higher in the case that interest rates have recently been high.<sup>78</sup>

Note that the same would be true of the output gap, if it were observed by the central bank and used as an indicator of inflationary pressures. In the optimal equilibrium, the output gap first increases in response to an increase in the natural rate (the direct effect  $x_n > 0$ ), but subsequently undershoots its long-run value as a result of the tightening of monetary policy ( $x_r < 0$ ). This would imply that if we seek to achieve the optimal equilibrium through any feedback rule with  $\phi_n = 0$ , and with  $\phi_\pi$  and  $\phi_x$  of the same sign, we would need to use a rule with  $\phi_\pi, \phi_x > 0$  (as proposed by Taylor), *and* with  $\theta > r_r$ . (The conclusion that  $\phi_\pi, \phi_x$  must both be positive if they have the same sign follows directly from (5.7), if the desired equilibrium involves  $\pi_n, x_n > 0$ . The conclusion that  $\theta > r_r$  then follows from (5.6),

---

<sup>78</sup>Note that this argument has nothing to do with the idea that the disturbances to the natural rate are themselves serially correlated, so that evidence of a high natural rate in the recent past leads one to infer a higher current natural rate. For even in the case that  $\rho = 0$ , we have seen that the optimal impulse responses involve an initial increase in inflation followed by subsequent undershooting, so that also in this case the optimal coefficients satisfy  $\theta > r_r > 0$ . In the present case, the current value of the natural rate can be perfectly inferred from observations of inflation, so there is no need to use estimates of the natural rate in the past to infer the current value. In the more realistic case, in which observations of current conditions do not suffice to fully reveal the state of the economy, the argument suggested above *can* be a further reason for an optimal interest-rate feedback rule to be inertial, as shown by Aoki (1998).

given that the desired equilibrium involves  $\pi_r, x_r < 0$ .) Thus the conclusion that a desirable feedback rule involves  $\theta > r_r$  does not depend upon the exclusive use of inflation itself as an indicator of underlying inflationary pressures.

In fact, the degree of inertia indicated by estimated central bank reaction functions is typically greater than the optimal degree of intrinsic inertia in the response to shocks computed above (an inertia coefficient  $r_r$  of less than .5, at the quarterly frequency, for our calibrated parameter values). But given that such reaction functions respond to endogenous indicators of inflationary pressure, such as inflation and output, rather than to exogenous disturbances directly, it is in fact optimal for them to involve inertia coefficients greater than the value of  $r_r$ .

Indeed, if anything, the degree of inertia implied by estimated reaction functions is *less* than would be indicated as optimal on the grounds considered here. This can be illustrated in an especially dramatic way by considering the properties of the equilibria associated with rules of the form (5.8), for alternative values of  $\phi_\pi$  and  $\theta$ . Figures 13-16 describe the numerical solution for the stationary rational expectations equilibrium associated with such a feedback rule, in the case of the parameter values given in Table 1, including the specification  $\rho = .35$ . In each of these figures, the set of coefficients for which rational expectations equilibrium is indeterminate is indicated by the grey region. Note that in the non-negative orthant (the region of these figures of primary interest to us), the condition for determinacy is given by (5.5), as noted above.<sup>79</sup>

In the zone of determinacy, the figures display contour plots for the statistics  $V[\pi]$ ,  $V[x]$ ,  $V[r]$ , and  $E[W]$ . One observes that from the point of view of stabilization of inflation and of the output gap alone, there is no advantage to rules with  $\theta \neq 0$ ; the greatest degree of stabilization occurs if  $\phi_\pi$  is made as large as possible, but in the case of a sufficiently large value of  $\phi_\pi$ , the variability of both target variables is minimized by choosing  $\theta = 0$ . However, a rule with  $\theta > 0$  results in less volatile short-term interest rates in equilibrium; in fact, for

---

<sup>79</sup>The same condition applies in the more complicated model of Rotemberg and Woodford (1998). See the corresponding figures in that paper.

given  $\phi_\pi > 0$ , the variability of the nominal interest rate is minimized by choosing  $\theta$  as large as possible. Minimization of our overall loss criterion  $E[W]$  involves a compromise between these two objectives; it is optimal to choose large positive (but finite) values for both  $\phi_\pi$  and  $\theta$ . For the parameter values used here, the loss-minimizing policy rule coefficients are  $\phi_\pi = 46.1$  and  $\theta = 13.0$ .

This value of  $\theta$  implies a great deal of inertia indeed in the rule used by the central bank to set interest rates. The optimal value turns out to be decreasing in the assumed coefficient of autocorrelation of the natural rate of process. But for any value of  $\rho$ , the optimal rule within the family (5.8) is a “super-inertial” one with  $\theta > 1$ .<sup>80</sup>

### 5.3 A Feedback Rule that Implements the Optimal Plan

We now turn briefly to the question of designing an interest-rate feedback rule that can implement exactly the optimal pattern of responses to shocks described in section 3. As above, there will be many feedback rules that equally share the property of having a determinate rational expectations equilibrium that corresponds to the optimal plan. Rather than trying to characterize all such rules, we shall restrict our attention here to rules that involve feedback only from the inflation rate, and from past values of the interest rate itself, *i.e.*, rules of the form

$$A(L)\hat{r}_t = B(L)\hat{\pi}_t, \tag{5.10}$$

where  $A(L)$  and  $B(L)$  are two finite-order polynomials in the lag operator. The degree to which such a rule involves intrinsic interest-rate inertia of the kind discussed in the introduction is indicated by the lag polynomial  $A(L)$ , and in particular by the roots of the characteristic polynomial

$$A(\mu^{-1}) = 0. \tag{5.11}$$

As in the simpler case just considered, there is no simple relation between the form of the lag polynomial  $A(L)$  in the feedback rule (5.10) and that of the lag polynomial  $Q(L)$  in

---

<sup>80</sup>See Giannoni (1999) for further discussion of how the optimal simple rule depends upon model parameters, and for an analysis of “robust” optimal policy taking account of parameter uncertainty. Giannoni finds that in the case of parameter uncertainty the robust-optimal simple rule involves an even *larger* value of  $\theta$ .

(3.32) describing how interest rates evolve in response to shocks, in the stationary rational expectations equilibrium associated with such a policy rule. In particular, a “super-inertial” policy, in which  $A(\mu^{-1})$  has a root larger than one, may nonetheless be consistent with a stationary rational expectations equilibrium, even though stationarity implies that all roots of  $Q(\mu^{-1})$  must lie within the unit circle.

We again restrict attention to the case of a natural rate process described by (3.15). In this case, the vector of endogenous variables  $q_t \equiv [\hat{r}_t \hat{\pi}_t]'$  that enter the monetary policy rule (5.10) evolve according to a law of motion of the form

$$q_t = H\phi_{t-1} + hr_t^n \quad (5.12)$$

under the optimal plan, while the Lagrange multipliers evolve according to

$$\phi_t = N\phi_{t-1} + nr_t^n. \quad (5.13)$$

Here  $N$  is the same matrix as in (3.29) and (3.31), while  $n$  is the column vector defined by

$$n \equiv -C \sum_{j=0}^{\infty} \rho^j \tilde{A}^{-(j+1)} a.$$

The first row of  $H$  is given by the row vector  $p'$  from (3.30), and the second row is given by the first row of the matrix  $G$  in (3.28). Finally, the column vector  $h$  is defined by

$$h \equiv - \begin{bmatrix} q' \\ e'_1 \end{bmatrix} \sum_{j=0}^{\infty} \rho^j \tilde{A}^{-(j+1)} a,$$

where  $q'$  is the row vector in (3.30), and we use  $e_i$  to denote the  $i$ th unit vector (*i.e.*, the vector whose  $i$ th element is one, and all other elements zero).

A first question is then whether there exists a relation of the form (5.10) such that this relation holds at all times if the vector  $q_t$  evolves according to (5.12) – (5.13). It can be shown that the answer is yes, and that the lowest-order polynomials<sup>81</sup> for which this is possible are a second-order polynomial  $A(L)$  and a first-order polynomial  $B(L)$ . It is furthermore shown

---

<sup>81</sup>Obviously, for any polynomials  $A(L)$  and  $B(L)$  with this property, the polynomials  $\tilde{A}(L) \equiv C(L)A(L)$  and  $\tilde{B}(L) \equiv C(L)B(L)$  will also have the property, where  $C(L)$  is any lag polynomial whatsoever. But the latter, higher-order polynomials represent simply a more complex description of an equivalent policy rule.

that there is a unique such relation with polynomials of this order; specifically, it is the unique such relation in which neither  $A(L)$  nor  $B(L)$  is of order greater than two. This relation can be written in the form

$$\sum_{j=0}^2 v'_j q_{t-j} = 0, \quad (5.14)$$

where the three vectors of coefficients  $v'_j$  are given by

$$\begin{aligned} v'_0 &= e'_1 R^{-1}, \\ v'_1 &= -e'_1 R^{-1} H P Q^{-1}, \\ v'_2 &= e'_1 R^{-1} H [P Q^{-1} - N H^{-1}] H N H^{-1}, \end{aligned}$$

and

$$P \equiv [n \quad Nn - N^2 H^{-1} h], \quad Q \equiv [h \quad Hn - H N H^{-1} h], \quad R \equiv [e_1 \quad h].$$

As one can show furthermore that this definition implies that  $v'_2 e_2 = 0$ , the polynomials  $A(L)$  and  $B(L)$  implied by (5.14) are indeed of the asserted order.

There is thus a relation of the desired form that is *consistent* with the optimal pattern of responses of inflation and interest rates to disturbances to the natural rate. It remains for us to determine whether imposition of (5.14) as a policy rule would result in a determinate equilibrium. The algebraic conditions that determine this are rather complex, but numerical investigation indicates that, at least in the case of parameters in the area of those listed in Table 1, the conditions for determinacy are satisfied by this rule. Thus we have found a policy rule that is not only consistent with the optimal equilibrium responses to shocks, but also satisfies both of our additional desiderata: its implementation does not require the central bank to observe the current values of the natural rates of interest and output, and it results in a determinate equilibrium.

What does the form of this optimal interest-rate feedback rule indicate about the desirability of interest-rate inertia of the kind indicated by the empirical studies summarized in the introduction? Our characterization here of the optimal form of the lag polynomial  $A(L)$  in (5.10) is arguably more relevant for that question than our previous characterization of

the optimal form of the lag polynomial  $Q(L)$  in (3.32), since a number of the studies referred to seek to estimate an explicit interest-rate feedback rule intended to represent systematic Fed policy. It is thus of particular interest to note that (5.14) implies that the coefficients on the lags in  $A(L)$  should again be substantial in magnitude.

Indeed, for plausible parameter values, these coefficients imply an even greater degree of interest-rate inertia than was suggested by the coefficients on the lags in  $Q(L)$ . Again, it is useful to consider the roots of the characteristic polynomial (5.11). For example, in the case of the parameter values given in Table 1, and our baseline natural-rate process with  $\rho = .35$ , the optimal rule (5.8) takes the form

$$A(L) = 1 - 12.9L + 8.3L^2 = (1 - 12.2L)(1 - .68L), \quad B(L) = 42.6 - 27.8L = 42.6(1 - .65L).$$

The roots of (5.11) are thus 12.2 and .68; the larger of these is very large indeed.

Furthermore, it is clearly the larger root, in this case, that mainly determines the degree of interest-rate inertia implied by the specification (5.10). For one observes that  $A(L)$  and  $B(L)$  are close to having a common factor  $(1 - .65L)$ ; thus, for these parameter values, (5.10) can be approximated by a rule of the simpler form (5.8), with coefficients  $\phi_\pi = 42.6$ ,  $\theta = 12.2$ . (Note that these coefficients are close to the ones found above to be optimal among rules of that class.) Thus, as concluded earlier, a simple rule of that form provides a good approximation to an optimal policy.

## 5.4 Discussion

This same conclusion – that the optimal policy rule involves an autoregressive polynomial with a largest root that is even greater than one – also obtains in the more complex model studied by Rotemberg and Woodford (1998). In that model, there are two independent types of innovations each period to the state vector  $s_t$  that describes current information about the future evolution of the *IS* and *AS* disturbances. (The two orthogonal shocks are identified in the U.S. data with two different orthogonal innovations from a tri-variate VAR.) An optimal plan involves interest-rate responses to each of these kinds of news, and as a

result feedback from the current inflation rate alone, along with lagged variables, does not allow a sufficiently flexible class of responses to include one that achieves the exact optimal plan. However, it is still possible to find a policy rule that supports the optimal plan, and that does not require the central bank to observe the current values of any of the exogenous shocks for its implementation. This rule is of the form

$$A(L)\hat{r}_t = B(L)\hat{\pi}_t + C(L)\hat{y}_t,$$

where  $\hat{y}_t$  represents the log of detrended real output, rather than the output gap. Rotemberg and Woodford find that for the parameter values associated with their estimated model, the characteristic equation (5.11) has a largest root  $\mu = 1.33$ ; this is less explosive than the optimal lag polynomial in the calculations just reported, but it is explosive nonetheless.

The presence of an autoregressive root greater than one means that in the case of arbitrary bounded fluctuations in inflation, interest-rate determination according to (5.10) would in general imply explosive fluctuations in the interest rate, that would not remain forever within any finite bounds. However, this does *not* mean that in the rational expectations equilibrium associated with such a rule, interest rates will be highly volatile. Instead, the stationary equilibrium associated with this rule involves a lower variance of the stationary distribution of nominal interest rates than is associated, for example, with the time-consistent optimizing policy, or with the optimal non-inertial policy, as shown in Table 2. This is because, in equilibrium, the anticipation that interest rates will be set according to (5.10) results in inflation fluctuations that keep nominal interest rates from exploding, or even from deviating very far from their steady-state level.

An unexpected inflationary shock (an unexpected increase in the natural rate of interest) requires some initial increase in the level of nominal interest rates. But it does not imply further explosive growth of interest rates in subsequent quarters, because inflation falls, and indeed undershoots its long-run level, after its initial small increase. Given this undershooting, the policy rule (5.10) allows the nominal interest rate to converge back to its normal level after a few quarters. The inflation reduction occurs, in equilibrium, because if demand

did *not* decline enough, and price increases were *not* moderated enough, to achieve this, the private sector would have to expect steep interest-rate increases a few quarters in the future, and that expectation would justify sharper reductions of demand and of price increases. The fact that the rule requires this to occur in equilibrium, of course, is exactly what makes it an optimal policy, since as we saw in section 3, an optimal plan would arrange for that to occur, in order to reconcile low inflation variability with low interest-rate variability.

Confidence that the private sector will in fact adopt spending and pricing behavior that keeps the economy on a non-explosive path, as in the stationary rational expectations equilibrium, depends, of course, upon the credibility of the central bank's commitment to the inertial rule. A consideration of what kind of rule would be optimal in the case that the rule is not expected to be perfectly credible, or if the private sector is expected to learn about the new regime only over time, is beyond the scope of the present study. However, it is worth noting that "super-inertial" rules need not produce such dramatically bad outcomes, even if the behavioral assumptions of our forward-looking model are incorrect, as is suggested by exercises like that of Rudebusch and Svensson (1998). These authors substitute a super-inertial interest-rate feedback rule, found to have desirable properties in the model of Rotemberg and Woodford (1998) for reasons essentially the same as those analyzed here, into a backward-looking structural model (not derived from optimizing behavior), and find that it results in infinite variances, and hence a worse outcome than would be associated with other, less inertial rules.

This might make it seem that a concern for "robustness" should make one avoid super-inertial rules, simply because they lead to extremely bad outcomes under some possible assumptions (Taylor, 1998). However, it is important to note that the infinite variances result only if the rule is taken to apply, with constant coefficients, *no matter how extreme* the levels of inflation and interest rates may have become. The desirable properties of the rule in the model considered here, by contrast, depend only upon a commitment to apply the rule in the case of fluctuations in inflation and interest rates over bounded intervals, the range over which these variables fluctuate in the stationary rational expectations equilibrium. One

could modify the rule outside these intervals (for example, specifying that interest rates will never fall below a certain floor, or rise above a certain ceiling, no matter what the history of inflation and interest rates may be), and still obtain the same prediction of a determinate rational expectations equilibrium with the desirable properties analyzed above. Thus, while the question of the robustness of our results to model uncertainty is certainly an important one,<sup>82</sup> it seems premature to judge that super-inertial rules *as such* can be excluded from consideration.

It is clear that the particular numerical coefficients derived above for an optimal feedback rule do not deserve much emphasis, even if the model were viewed as a reasonable approximation of the U.S. economy. For these coefficients – for example, the optimal  $(\phi_\pi, \theta)$  in a rule of the form (5.8) – are quite sensitive to relatively small changes in the specification of the model. This is shown by a comparison of our results here with those of Rotemberg and Woodford (1998) for the same class of policy rules, despite the fact that we have “calibrated” the simpler model used here to resemble theirs to the extent possible. One can also show that even within the context of the model as specified here, the optimal values of  $\phi_\pi$  and  $\theta$  are quite sensitive to variation in the assumed values of structural parameters such as  $\kappa$  and  $\sigma$ .

The reason for this fragility is not too hard to see. In section 5.2 above, we have shown how a rule of the form (5.8) may be consistent with equilibrium interest-rate responses of the desired sort, which involve partial adjustment of the short nominal rate toward an increasing function of the current natural rate. But this depends, essentially, upon responding to current inflation and past interest rates in a way that responds appropriately to the level of the current natural rate of interest that can be *inferred* from those endogenous variables. The problem with such a strategy for monetary policy is that, in a near-optimal equilibrium, inflation does *not* respond much to an innovation in the natural rate (Figure 4), and so it is a poor indicator of the underlying pressures that monetary policy should respond to.<sup>83</sup> Because

---

<sup>82</sup>See Hansen and Sargent (1999) for an analysis of this question in the case of one particular description of the possible specification error.

<sup>83</sup>The problem is related to the problem discussed in Bernanke and Woodford (1997), that arises when

inflation increases by only a small amount in response to such a shock, the optimal feedback rule must respond to observed inflation variations with an extremely large coefficient. And, because it is necessary to correct for the amount that one would have expected inflation to decline in the absence of such a shock, which is an increasing function of the lagged interest rate, the lagged interest rate must receive an extremely large coefficient as well.

This is nonetheless a desirable feedback rule, under the assumption that the signal contained in variations in current inflation, while weak, is *perfectly observed*, so that one need only amplify the signal by responding very strongly to it. Indeed, in the absence of other indicators of changes in the natural rate of interest, it is the best that one can do. But it is clearly unrealistic, in practice, to assume that the relevant inflation variable is observed with complete precision, especially in real time.<sup>84</sup> At the same time, other relevant information is available, including estimates of the current output gap (the other relevant variable in our simple model), that, while possibly subject to greater uncertainty than estimates of current inflation, surely contain additional information. We do not here take up the question of how the optimal feedback rule is affected by taking account of such noise in the information available to the central bank.<sup>85</sup> But it seems likely that under a more realistic specification of available information, less weight would be put on the recent rate of inflation, and more on other indicators that, like the output gap in our simulations above, respond more to innovations in the natural rate in a desirable equilibrium. This is likely to considerably reduce the desired size of the inertia coefficient  $\theta$ .

Nonetheless, our basic argument for a value of  $\theta$  in excess of  $r_r$  would seem likely to extend to the more general setting. For this depends solely upon the fact that the endogenous variable that is used as an indicator of exogenous inflationary pressures will tend to first rise in response to such pressures, but later fall as a result of the tightening of monetary policy that occurs in a desirable equilibrium. Thus the degree of policy inertia implied by estimated

---

an inflation forecast is used as an indicator of the real disturbances, despite the fact that successful policy should large stabilize forecastable inflation.

<sup>84</sup>Orphanides (1997) documents the degree of error contained in real-time estimates of the current rate of inflation available to the Federal Open Market Committee in its deliberations.

<sup>85</sup>See Aoki (1998) for a treatment of related issues in a forward-looking model similar to that used here.

Fed reaction functions like that of Sack (1998b) seems unlikely to be greater than can be justified as part of an optimal policy commitment. Indeed, it is possible that a more detailed analysis would still conclude, like the simple treatment above, that current U.S. policy is not even inertial *enough*. Our finding that *even* a “super-inertial” autoregressive polynomial may be optimal shows how wrong it is to assume that evidence of inertial behavior provides *prima facie* evidence of incompetence.

# References

- Aoki, Kosuke, "On the Optimal Monetary Policy Response to Noisy Indicators," unpublished, Princeton University, November 1998.
- Backus, David, and John Driffill, "The Consistency of Optimal Policy in Stochastic Rational Expectations Models," CEPR discussion paper no. 124, August 1986.
- Barro, Robert J., and David B. Gordon, "A Positive Theory of Monetary Policy in a Natural Rate Model," *Journal of Political Economy* 91: 589-610 (1983).
- Bernanke, Ben S., and Michael Woodford, "Inflation Forecasts and Monetary Policy," *Journal of Money, Credit, and Banking*, 24: 653-84 (1997).
- Blanchard, Olivier J., and Charles Kahn, "The Solution of Linear Difference Equations under Rational Expectations," *Econometrica* 48: 1305-1311 (1980).
- Blinder, Alan S., *Central Banking in Theory and Practice*, Cambridge: M.I.T. Press, 1998.
- Calvo, Guillermo, "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12: 383-98 (1983).
- Caplin, Andrew, and John Leahy, "Monetary Policy as a Process of Search," *American Economic Review* 86: 689-702 (1996).
- Christiano, Lawrence J., and Christopher J. Gust, "Taylor Rules in a Limited Participation Model," NBER working paper no. 7017, March 1999.
- Clarida, Richard, Jordi Galí and Mark Gertler, "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," NBER working paper no. 6442, March 1998a.
- —, — — and — —, "Monetary Policy Rules in Practice: Some International Evidence," *European Economic Review* 42:1033-1068 (1998b).
- —, — — and — —, "The Science of Monetary Policy," *Journal of Economic Literature*, forthcoming 1999.
- Cook, Timothy, and Thomas Hahn, "The Effect of Changes in the Federal Funds Rate Target on Market Interest Rates in the 1970s," *Journal of Monetary Economics* 24: 331-352 (1989).

- Currie, David, and Paul Levine, *Rules, Reputation and Macroeconomic Policy Coordination*, Cambridge: Cambridge University Press, 1993.
- Fair, Ray C., "The Sensitivity of Fiscal Policy Effects to Assumptions about the Behavior of the Federal Reserve," *Econometrica* 46: 1165-1179 (1978).
- —, "An Analysis of a Macro-Econometric Model with Rational Expectations in the Bond and Stock Markets," *American Economic Review* 69: 539-552 (1979).
- —, and E. Philip Howrey, "Evaluating Alternative Monetary Policy Rules," *Journal of Monetary Economics* 38: 173-193 (1996).
- Friedman, Milton, "The Role of Monetary Policy," *American Economic Review* 58: 1-17 (1968).
- —, "The Optimum Quantity of Money," in *The Optimum Quantity of Money and Other Essays*, Chicago: Aldine, 1969.
- Fuhrer, Jeffrey C. and Geoffrey Moore, "Monetary Policy Trade-offs and the Correlation between Nominal Interest Rates and Real Output," *American Economic Review* 85: 219-239 (1995).
- Gali, Jordi, and Mark Gertler, "Inflation Dynamics: A Structural Econometric Analysis," unpublished, New York University, August 1998.
- Giannoni, Marc, "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward-Looking Model," unpublished, Princeton University, June 1999.
- Goodfriend, Marvin, "Interest Rate Smoothing in the Conduct of Monetary Policy," *Carnegie-Rochester Conference Series on Public Policy*, Spring 1991, 7-30.
- Goodhart, Charles A. E., "Why Do the Monetary Authorities Smooth Interest Rates?" Special Paper no. 81, LSE Financial Markets Group, February 1996.
- —, "Central Bankers and Uncertainty," Keynes Lecture, British Academy, October 29, 1998.
- Guthrie, Graeme, and Julian Wright, "The Optimal Form of Interest-Rate Targeting," unpublished, University of Auckland, December 1998.
- Hansen, Lars P., Dennis Epple, and William Roberds, "Linear-Quadratic Duopoly Models of Resource Depletion," in T.J. Sargent, ed., *Energy, Foresight and Strategy*, Washington: Resources for the Future, 1985.

- Hansen, Lars P., and Thomas J. Sargent, "Robustness and Commitment: A Monetary Policy Example," unpublished, Stanford University, March 1999.
- Kerr, William, and Robert G. King, "Limits on Interest Rate Rules in the IS Model," *Economic Quarterly*, Federal Reserve Bank of Richmond, Spring 1996, pp. 47-76.
- King, Robert G., and Alexander L. Wolman, "Inflation Targeting in a St. Louis Model of the 21st Century," *Federal Reserve Bank of St. Louis Review* 78(3): 83-107 (1996).
- — and — —, "What Should the Monetary Authority do When Prices are Sticky?" unpublished, University of Virginia, April 1998.
- Kiley, Michael T., "Monetary Policy under Neoclassical and New-Keynesian Phillips Curves, with an Application to Price Level and Inflation Targeting," FEDS discussion paper no. 1998-27, May 1998.
- Kydland, Finn E., and Edward C. Prescott, "Rules Rather than Discretion: The Inconsistency of Optimal Plans," *Journal of Political Economy* 85: 473-491 (1977).
- — and — —, "Dynamic Optimal Taxation, Rational Expectations and Optimal Control," *Journal of Economic Dynamics and Control* 2: 79-91 (1980).
- Levin, Andrew, Volker Wieland, and John C. Williams, "Robustness of Simple Monetary Policy Rules under Model Uncertainty," NBER working paper no. 6570, May 1998.
- Levine, Paul, "Should Rules be Simple?" CEPR discussion paper no. 515, March 1991. [Reprinted as chapter 6 of Currie and Levine (1993).]
- Lucas, Robert E., Jr., "Econometric Policy Evaluation: A Critique," *Carnegie-Rochester Conference Series on Public Policy*, 1, 1976, 19-46.
- Mankiw, N. Gregory, "The Optimal Collection of Seignorage: Theory and Evidence," *Journal of Monetary Economics* 20: 327-341 (1987).
- McCallum, Bennett T., "Price Level Determinacy with an Interest Rate Policy Rule and Rational Expectations," *Journal of Monetary Economics* 8: 319-329 (1981).
- —, "Issues in the Design of Monetary Policy Rules," in J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics*, Amsterdam: North-Holland, forthcoming 1999.
- —, and Edward Nelson, "An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis," unpublished, Carnegie-Mellon University, October 1997.

- — and — —, “Performance of Operational Policy Rules in an Estimated Semi-Classical Structural Model,” NBER working paper no. 6599, June 1998.
- Orphanides, Athanasios, “Monetary Policy Rules Based on Real-Time Data,” FEDS discussion paper no. 1998-3, Federal Reserve Board, December 1997.
- Persson, Torsten, and Guido Tabellini, “Introduction,” in *Monetary and Fiscal Policy, Volume 1: Credibility*, Cambridge: M.I.T. Press, 1994.
- Roberts, John M., “New Keynesian Economics and the Phillips Curve,” *Journal of Money, Credit and Banking* 27: 975-984 (1995).
- —, “Inflation Expectations and the Transmission of Monetary Policy,” FEDS discussion paper no. 1998-43, Federal Reserve Board, October 1998.
- Rogoff, Kenneth, “The Optimal Degree of Commitment to an Intermediate Monetary Target,” *Quarterly Journal of Economics* 100: 1169-1190 (1985).
- Rotemberg, Julio J., and Michael Woodford, “An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy,” *NBER Macroeconomics Annual 1997*, 297-346. [Expanded version circulated as NBER Technical Working Paper no. 233, May 1998.]
- —, and — —, “Interest-Rate Rules in an Estimated Sticky-Price Model,” NBER working paper no. 6618, June 1998.
- Rudebusch, Glenn D., “Federal Reserve Interest Rate Targeting, Rational Expectations and the Term structure”, *Journal of Monetary Economics*, 35, 1995, 245-74.
- —, and Lars E.O. Svensson, “Policy Rules for Inflation Targeting,” in John B. Taylor, ed., *Monetary Policy Rules*, Chicago: University of Chicago Press, forthcoming.
- Sack, Brian, “Does the Fed Act Gradually? A VAR analysis”, FEDS discussion paper no. 1998-17, Federal Reserve Board, March 1998a.
- —, “Uncertainty, Learning, and Gradual Monetary Policy,” FEDS discussion paper no. 1998-34, Federal Reserve Board, July 1998b.
- Sargent, Thomas J., *Macroeconomic Theory*, 2d ed., New York: Academic Press, 1987.
- —, and Neil Wallace, “Rational Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule,” *Journal of Political Economy* 83: 241-254 (1975).

- Sbordone, Argia M., "Prices and Unit Labor Costs: A New Test of Sticky Prices," Seminar Paper no. 653, Institute for International Economic Studies, Stockholm, October 1998.
- Schmitt-Grohé, Stephanie, and Martin Uribe, "Price Level Determinacy and Monetary Policy under a Balanced-Budget Requirement," *Journal of Monetary Economics*, forthcoming 1999.
- Soderlind, Paul, "Solution and Estimation of RE Macromodels with Optimal Policy," unpublished, Stockholm School of Economics, December 1998.
- Summers, Lawrence, "How Should Long Term Monetary Policy Be Determined?" *Journal of Money, Credit and Banking*, 23, August 1991, 625-31.
- Svensson, Lars E.O., "Price Level Targeting versus Inflation Targeting: A Free Lunch?" NBER working paper no. 5719, August 1996.
- —, "Inflation Targeting as a Monetary Policy Rule," *Journal of Monetary Economics*, forthcoming 1999.
- Taylor, John B., "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy* 39: 195-214 (1993).
- —, "The Robustness and Efficiency of Monetary Policy Rules as Guidelines for Interest Rate Setting by the European Central Bank," IIES seminar paper no. 649, August 1998.
- Walsh, Carl, "Optimal Contracts for Central Bankers," *American Economic Review* 85: 150-167 (1995).
- Watson, Mark W., "Explaining the Increased Variability in Long Term Interest Rates," unpublished, Princeton University, May 1999.
- Woodford, Michael, "The Optimum Quantity of Money," in B.M. Friedman and F.H. Hahn, eds. *Handbook of Monetary Economics*, vol. II, Amsterdam: North-Holland, 1990.
- —, "Control of the Public Debt: A Requirement for Price Stability?" NBER working paper no. 5684, July 1996.
- —, "Price-Level Determination under Interest-Rate Rules," unpublished, Princeton University, April 1999a.
- —, "Inflation Stabilization and Welfare," unpublished, Princeton University, June

1999b.

Yun, Tack, “Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles,”  
*Journal of Monetary Economics* 37: 345-370 (1996).

Table 1: “Calibrated” parameter values.

Structural parameters	
$\beta$	.99
$\sigma$	.157
$\kappa$	.024
Shock process	
$\rho$	.35
$\text{sd}(r^n)$	3.72
Loss function	
$\lambda_x$	.048
$\lambda_r$	.236

Table 2: Statistics for alternative policies.

Case:  $\rho = 0$ .

Policy	V[ $\pi$ ]	V[ $x$ ]	V[ $r$ ]	E[ $W$ ]
Time Consist.	.122	13.43	2.004	1.244
Non-Inertial	.122	13.43	2.004	1.244
Optimal	.070	9.76	.983	.774

Case:  $\rho = .35$ .

Policy	V[ $\pi$ ]	V[ $x$ ]	V[ $r$ ]	E[ $W$ ]
Time Consist.	.487	22.95	4.023	2.547
Non-Inertial	.211	9.92	6.720	2.279
Optimal	.130	10.60	1.921	1.097

Case:  $\rho = .9$ .

Policy	V[ $\pi$ ]	V[ $x$ ]	V[ $r$ ]	E[ $W$ ]
Time Consist.	402.9	528.2	413.7	526.3
Non-Inertial	.353	.463	10.41	2.836
Optimal	.400	4.74	6.77	2.228

Table 3: Time-Consistent Equilibria with Alternative Policy Weights.

Policy Weights				Equilibrium Statistics			
$\hat{\lambda}_\pi$	$\hat{\lambda}_x$	$\hat{\lambda}_r$	$\lambda_\Delta$	$V[\pi]$	$V[x]$	$V[r]$	$E[W]$
1	.048	.236	0	.487	22.95	4.023	2.547
1	.048	.120	0	.211	9.92	6.720	2.279
1	.048	0	.282	.082	11.73	2.907	1.337
1	.048	-.439	.807	.135	10.50	1.922	1.097
1	0	-.232	.204	.135	10.50	1.922	1.097
20.7	1	-9.09	16.7	.135	10.50	1.922	1.097
0	1	0	5.51	.077	11.95	2.649	1.281

Natural, Optimal, and Optimal Non-inertial Interest Rate [ $\rho = 0$ ]

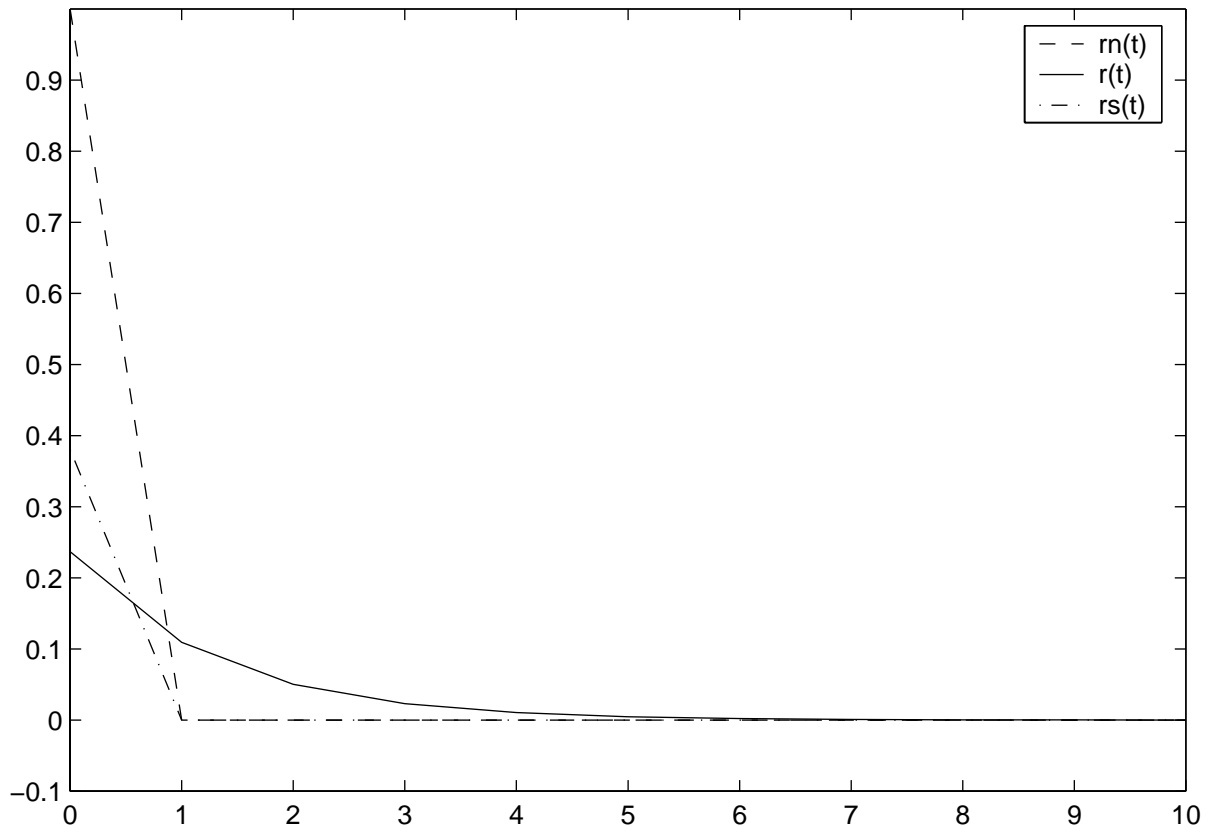


Figure 1

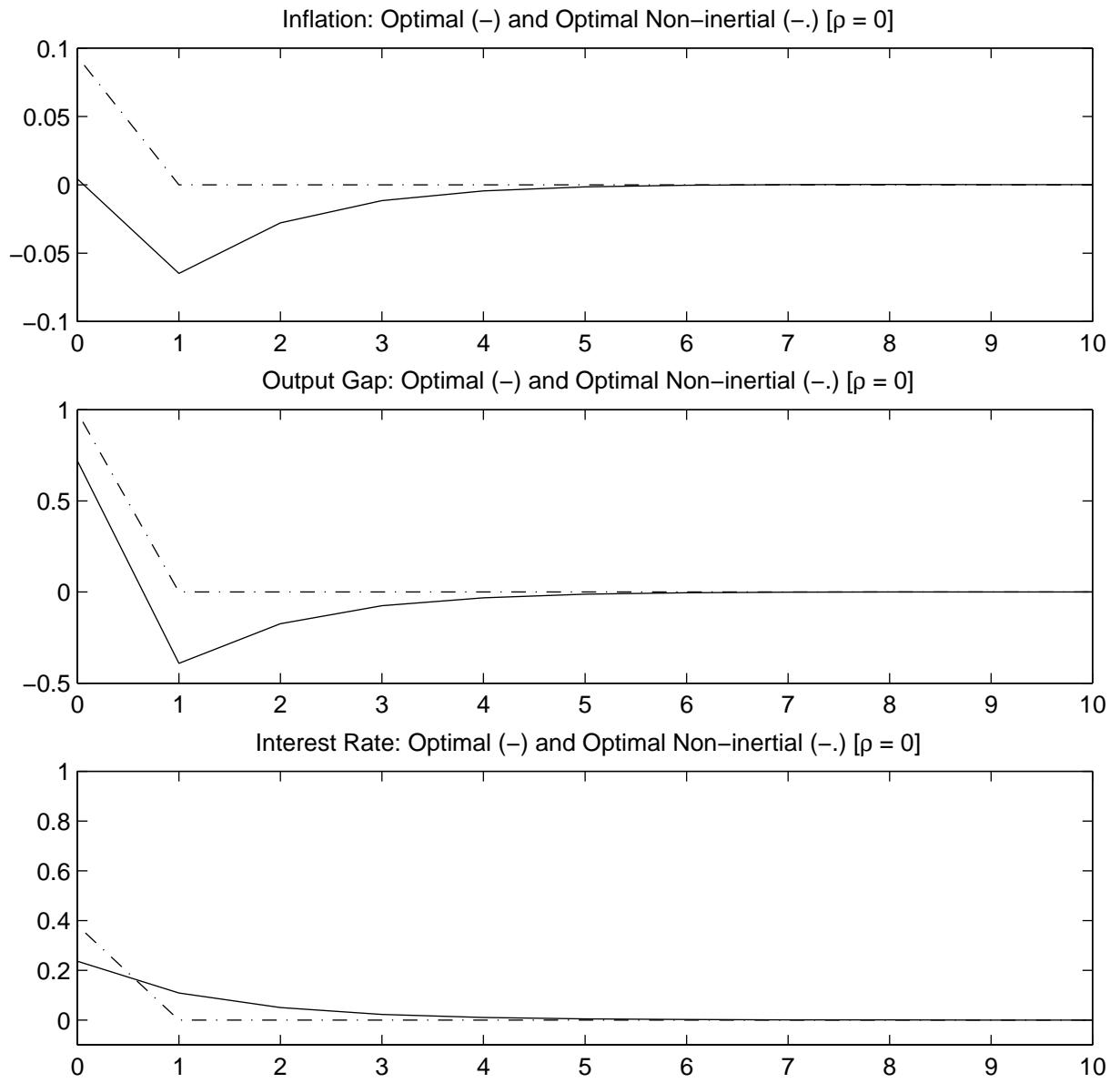


Figure 2

Natural, Optimal, and Optimal Non-inertial Interest Rate [ $\rho = 0.35$ ]

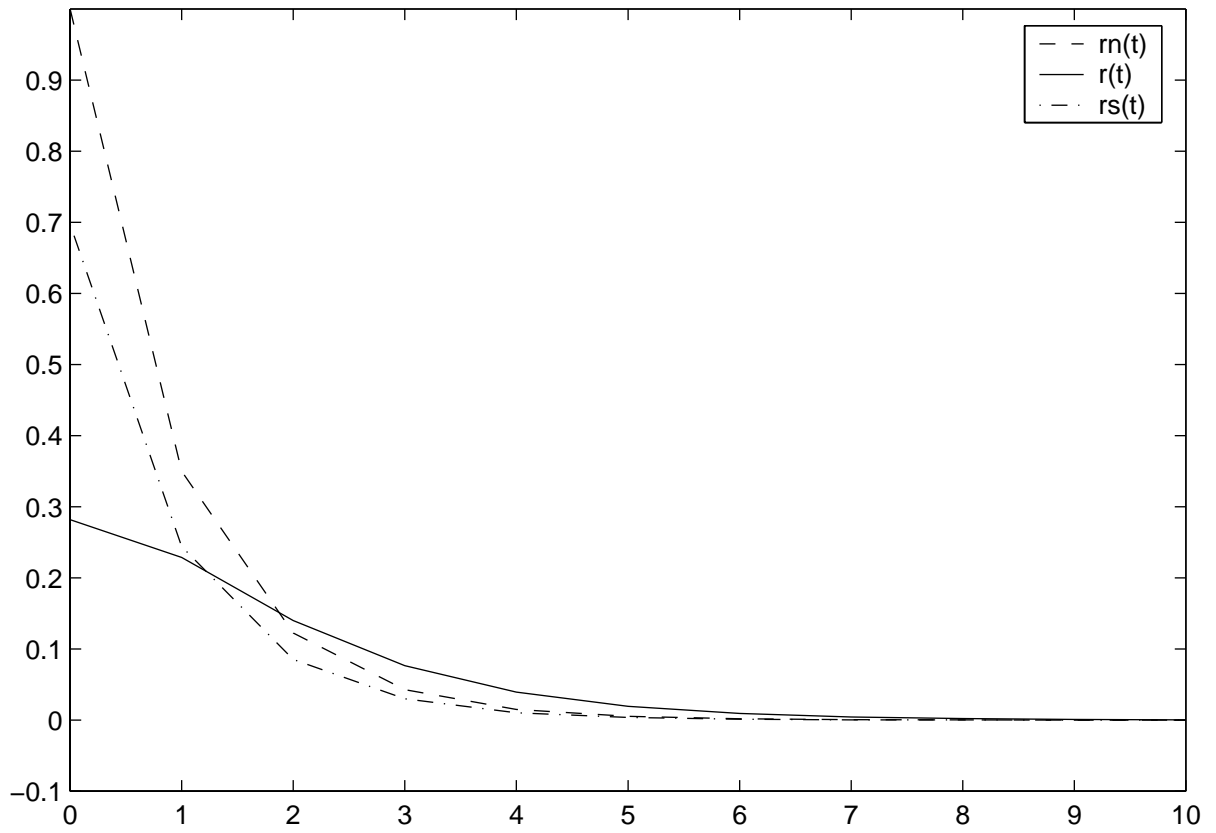


Figure 3

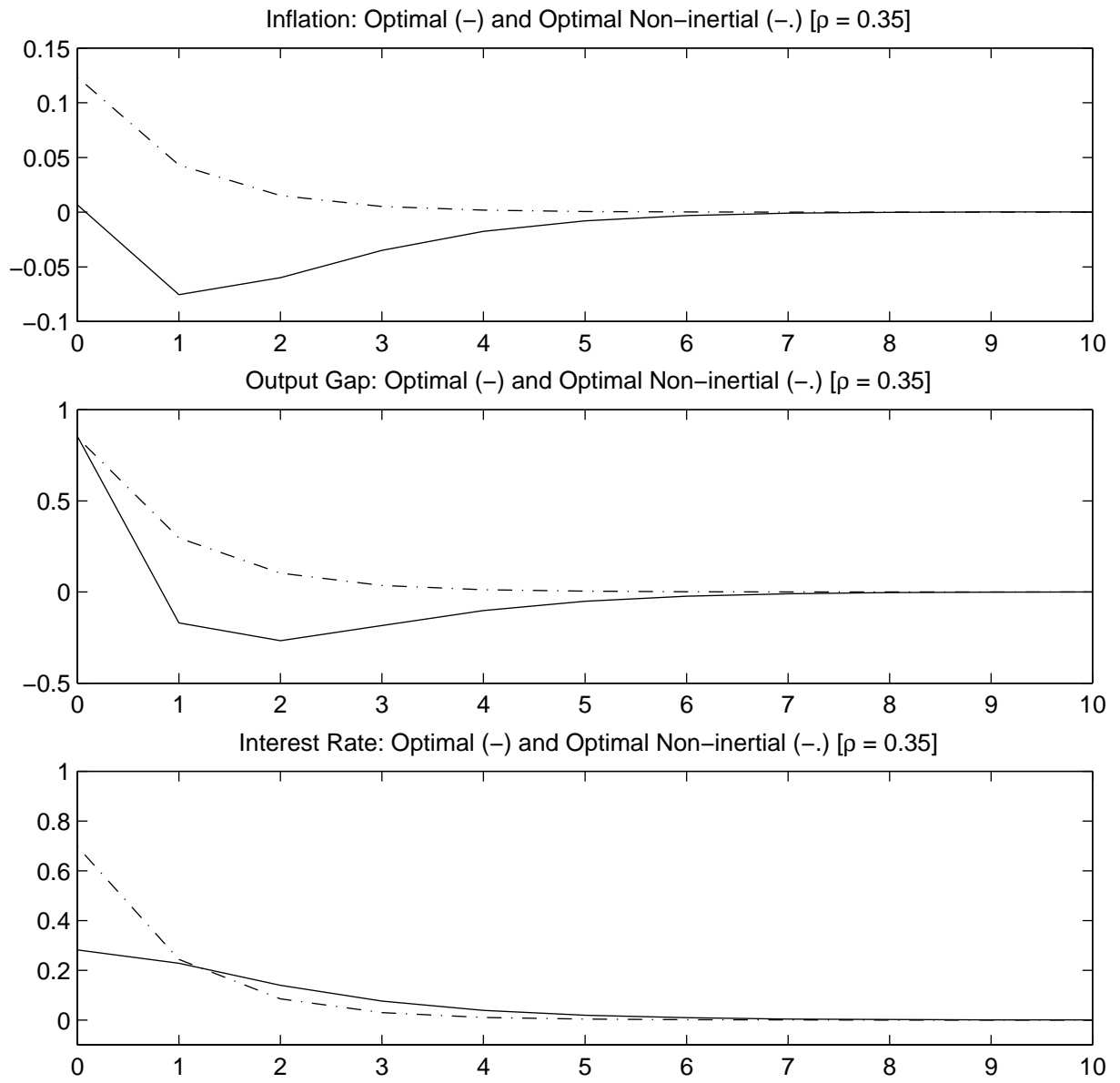


Figure 4

Natural, Optimal, and Optimal Non-inertial Interest Rate [ $\rho = 0.9$ ]

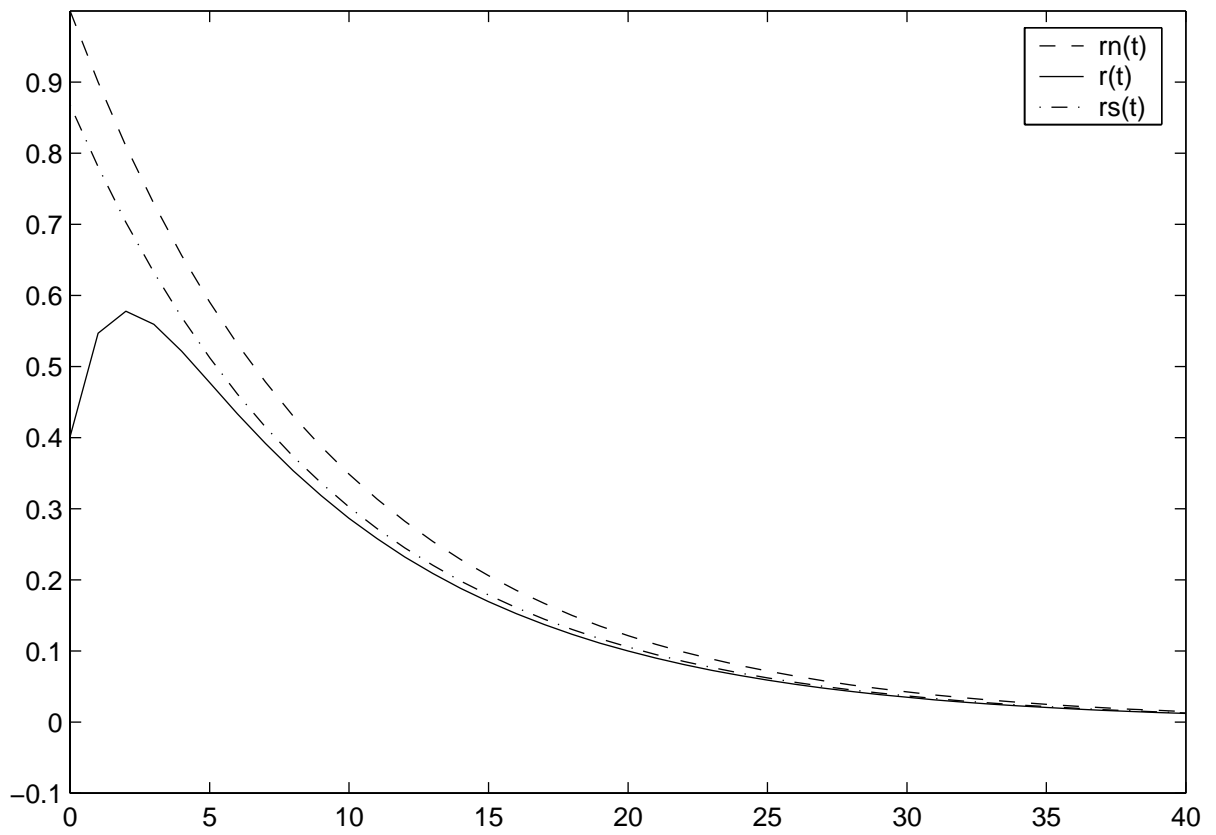


Figure 5

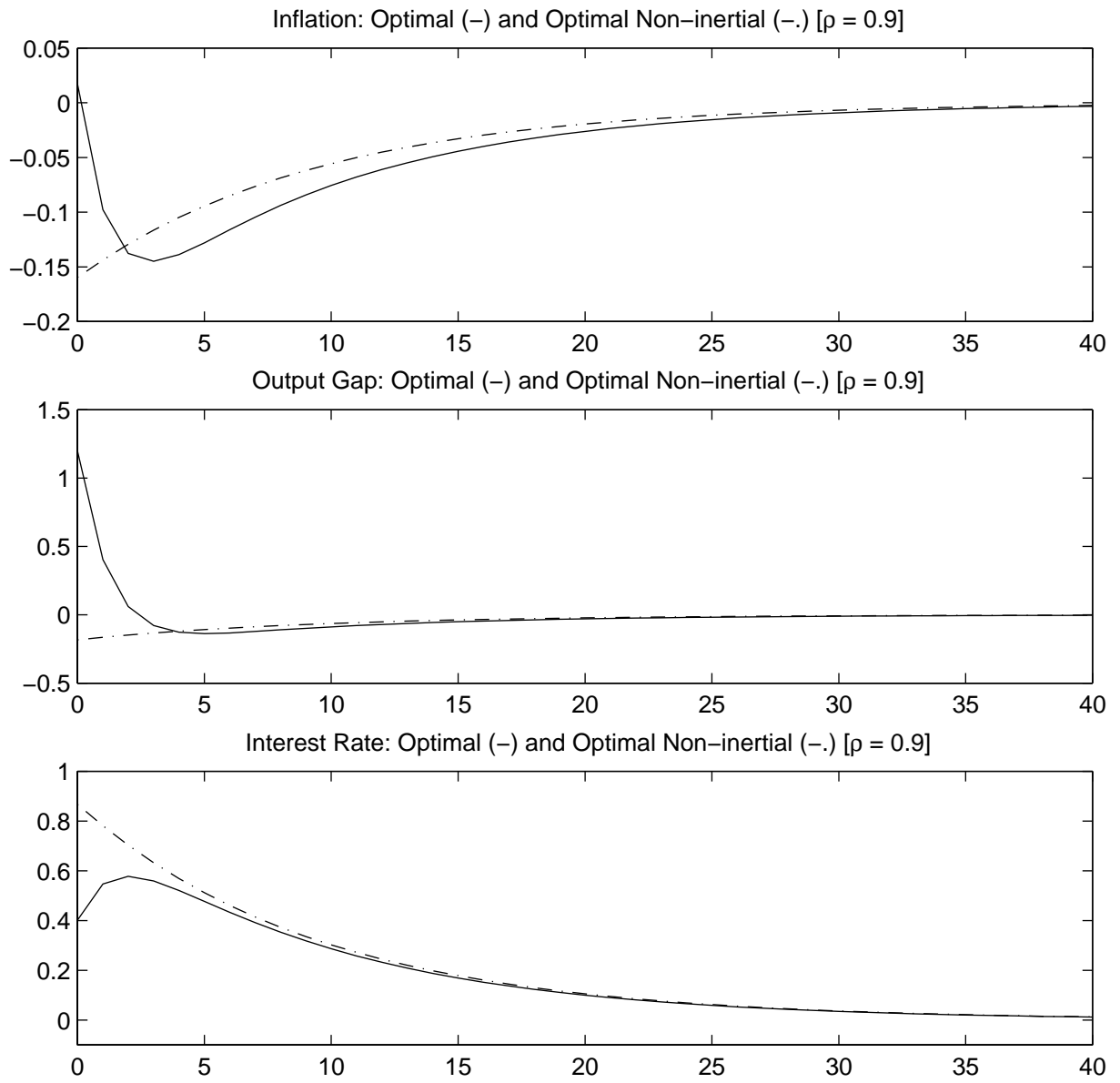


Figure 6

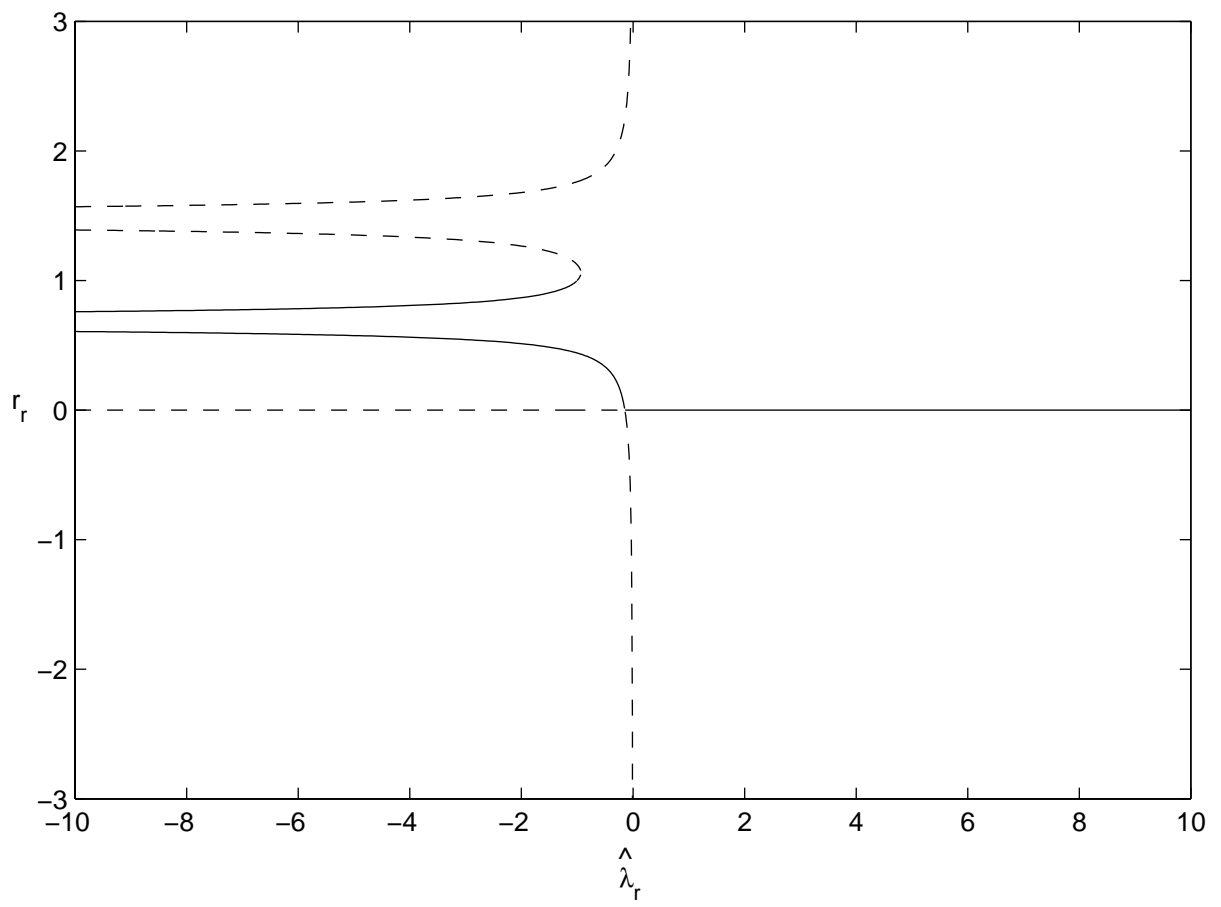


Figure 7

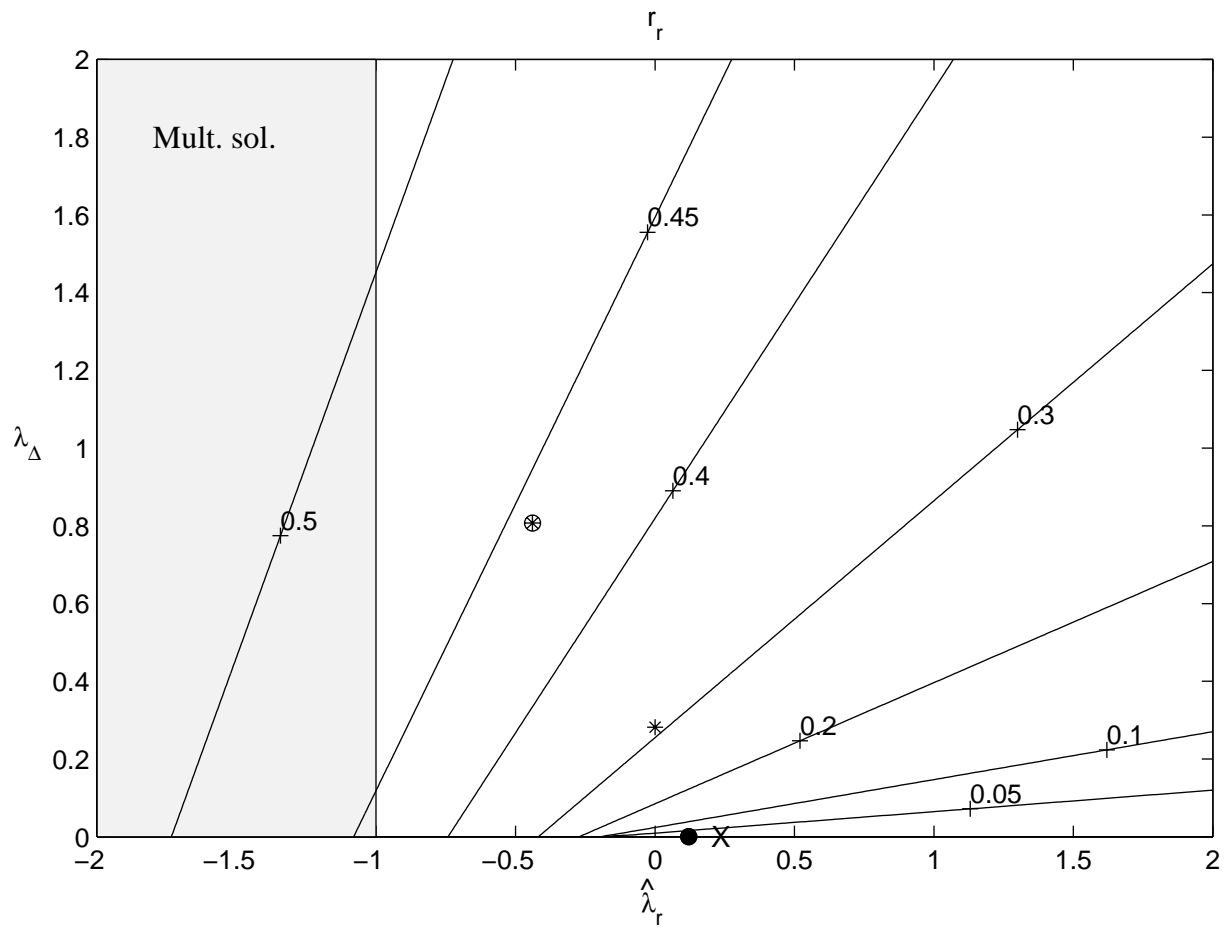


Figure 8

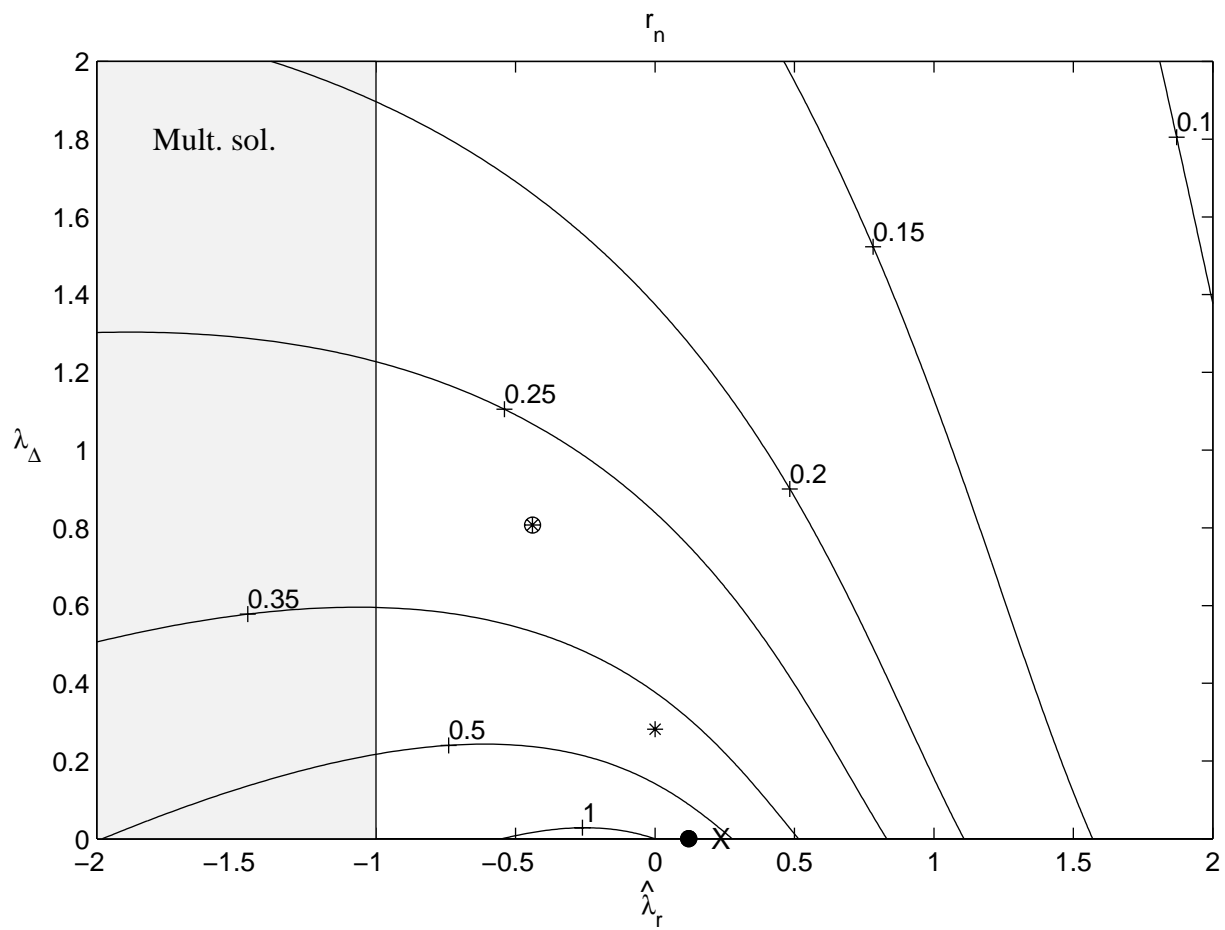


Figure 9

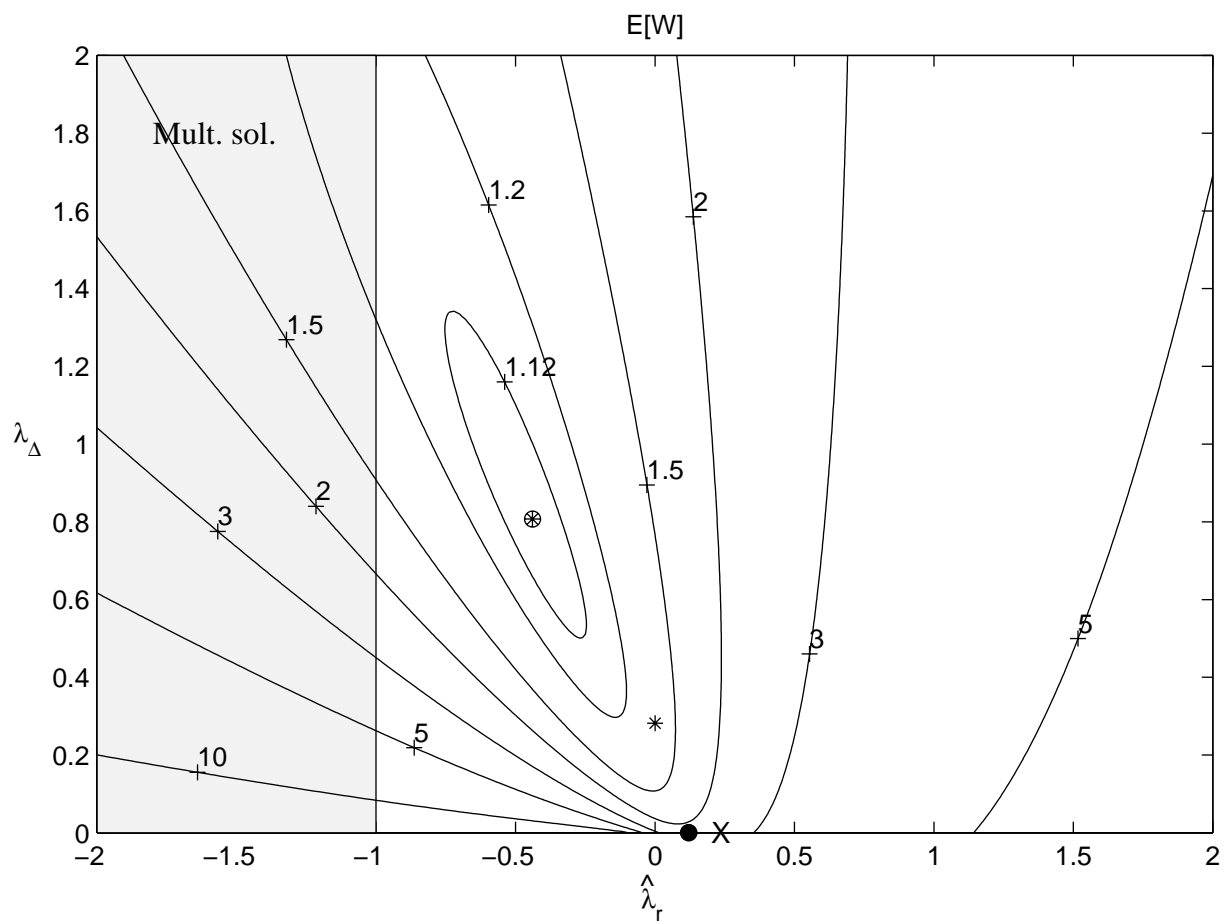


Figure 10

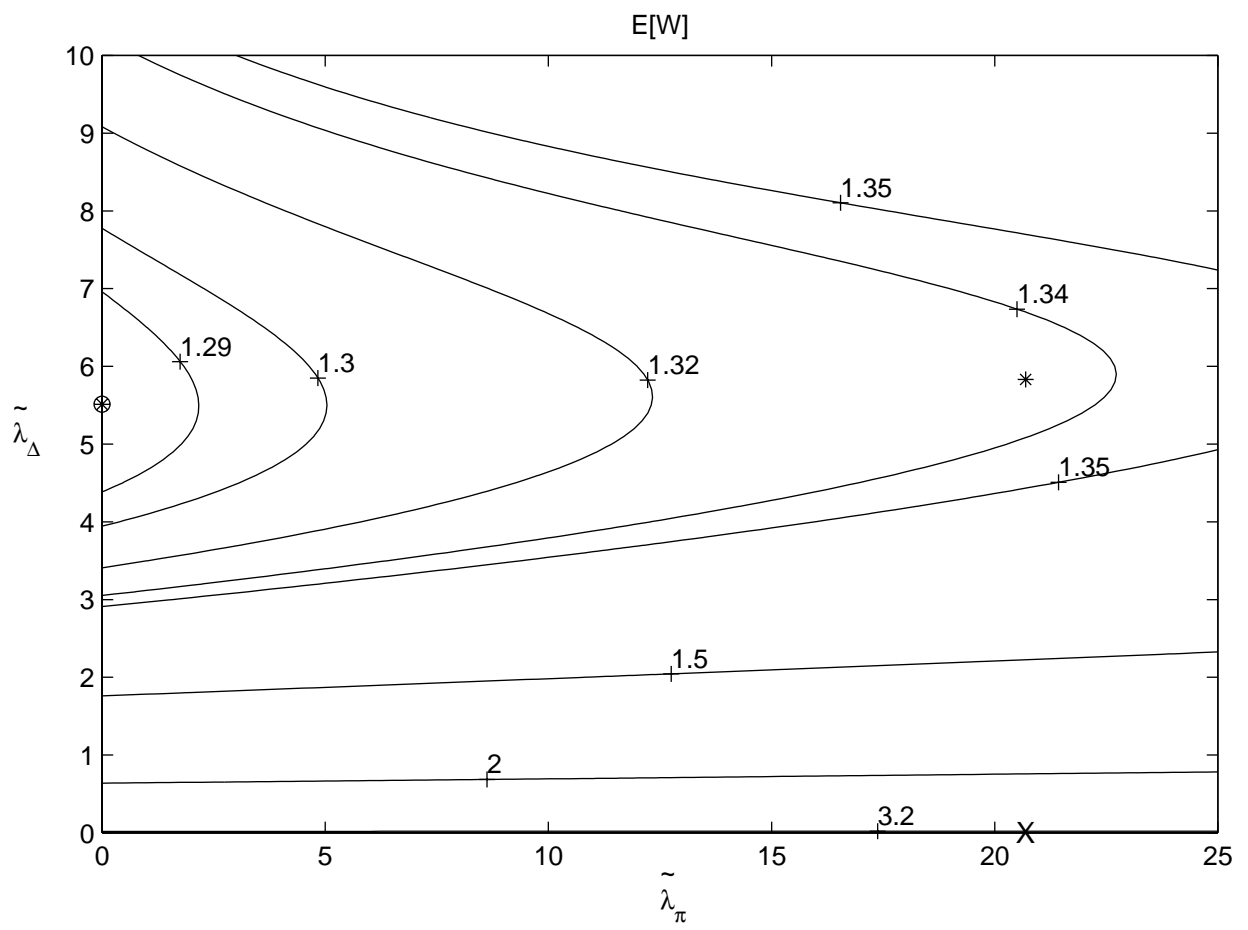


Figure 11

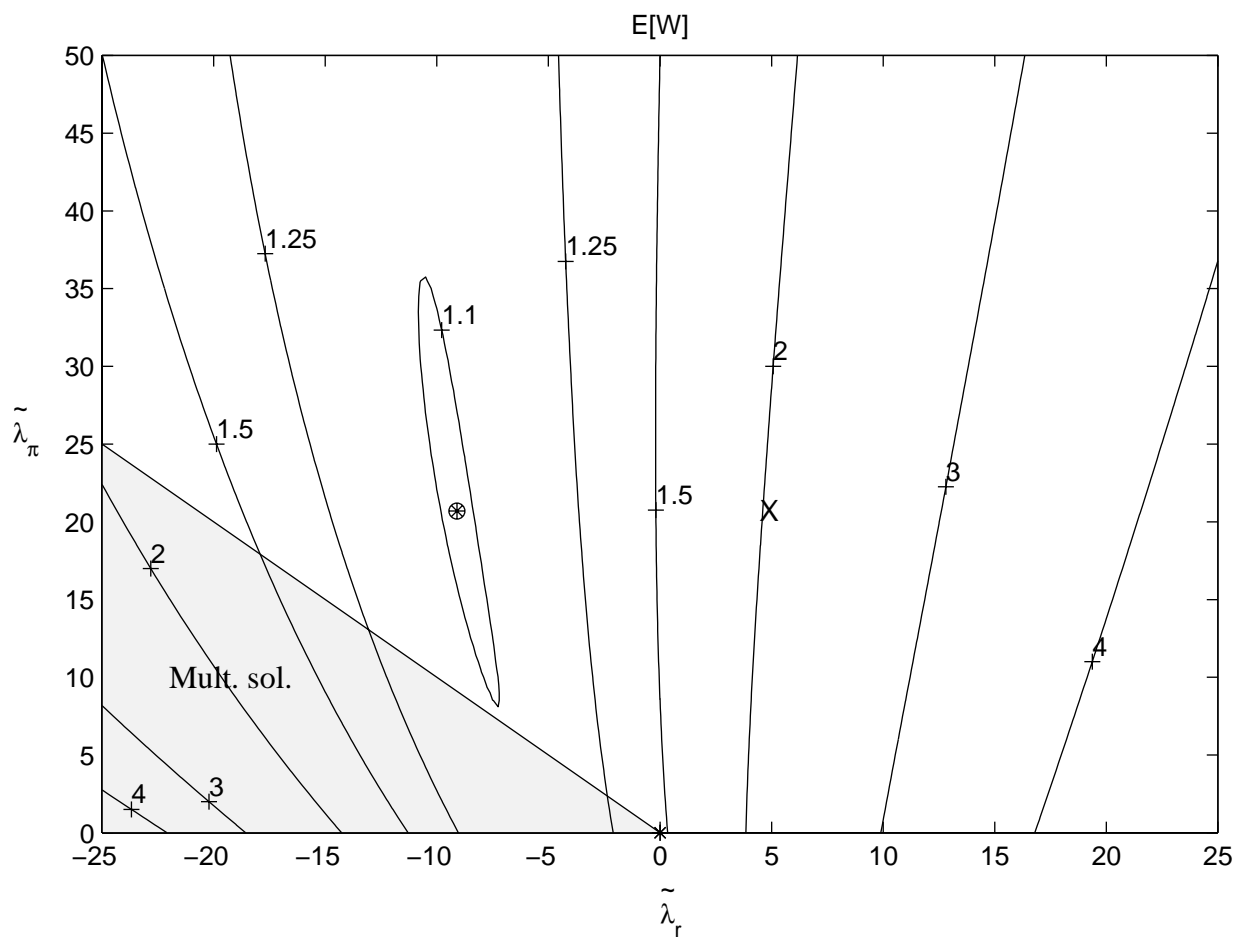


Figure 12

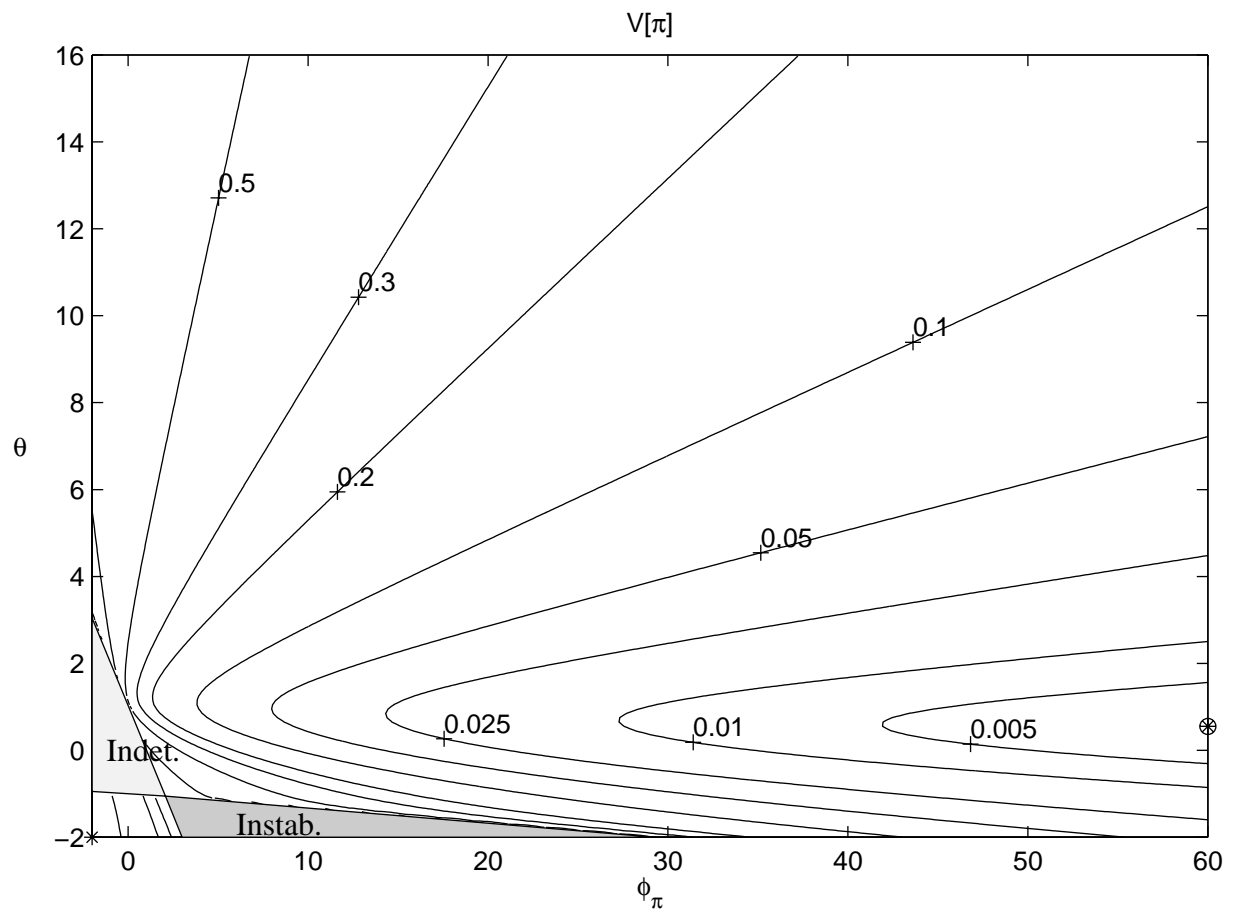


Figure 13

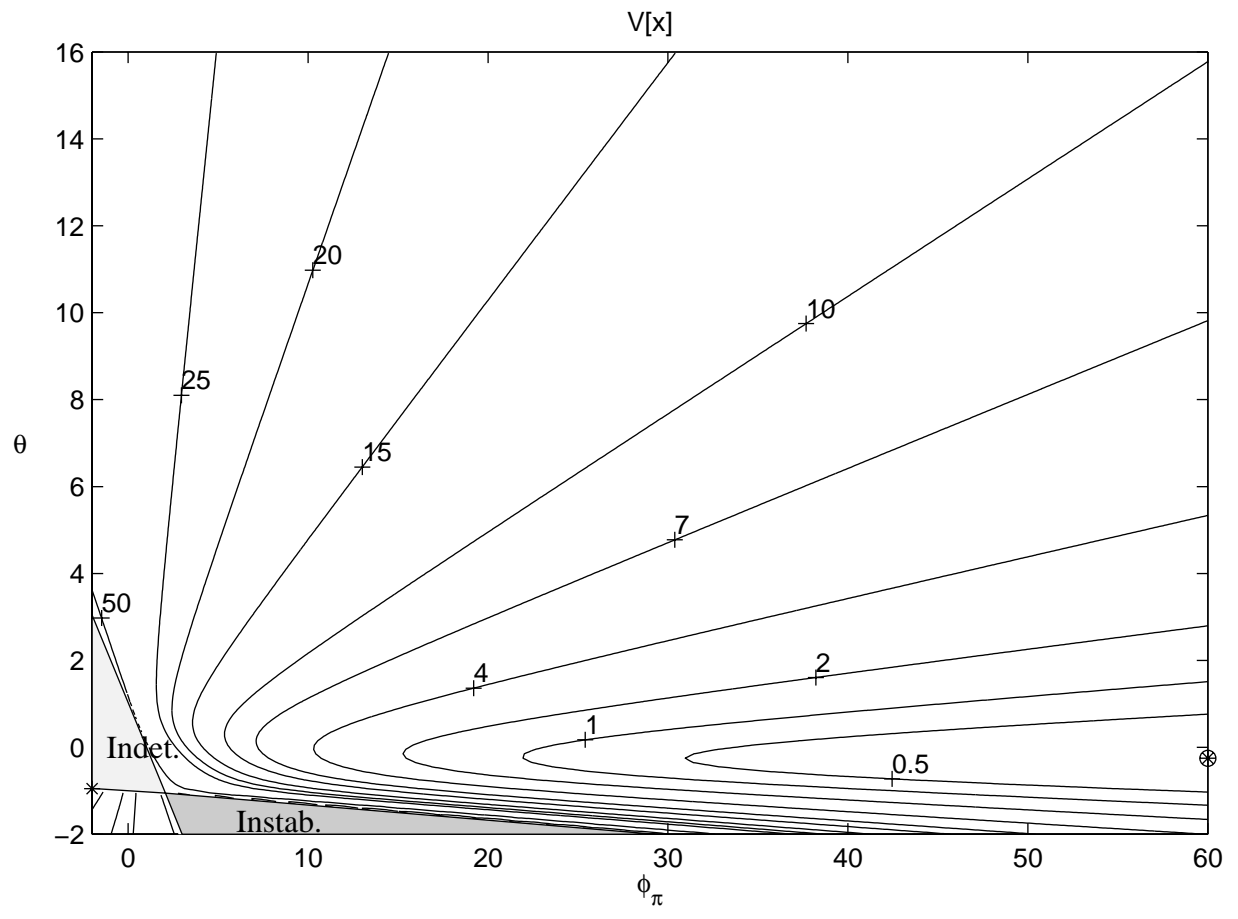


Figure 14

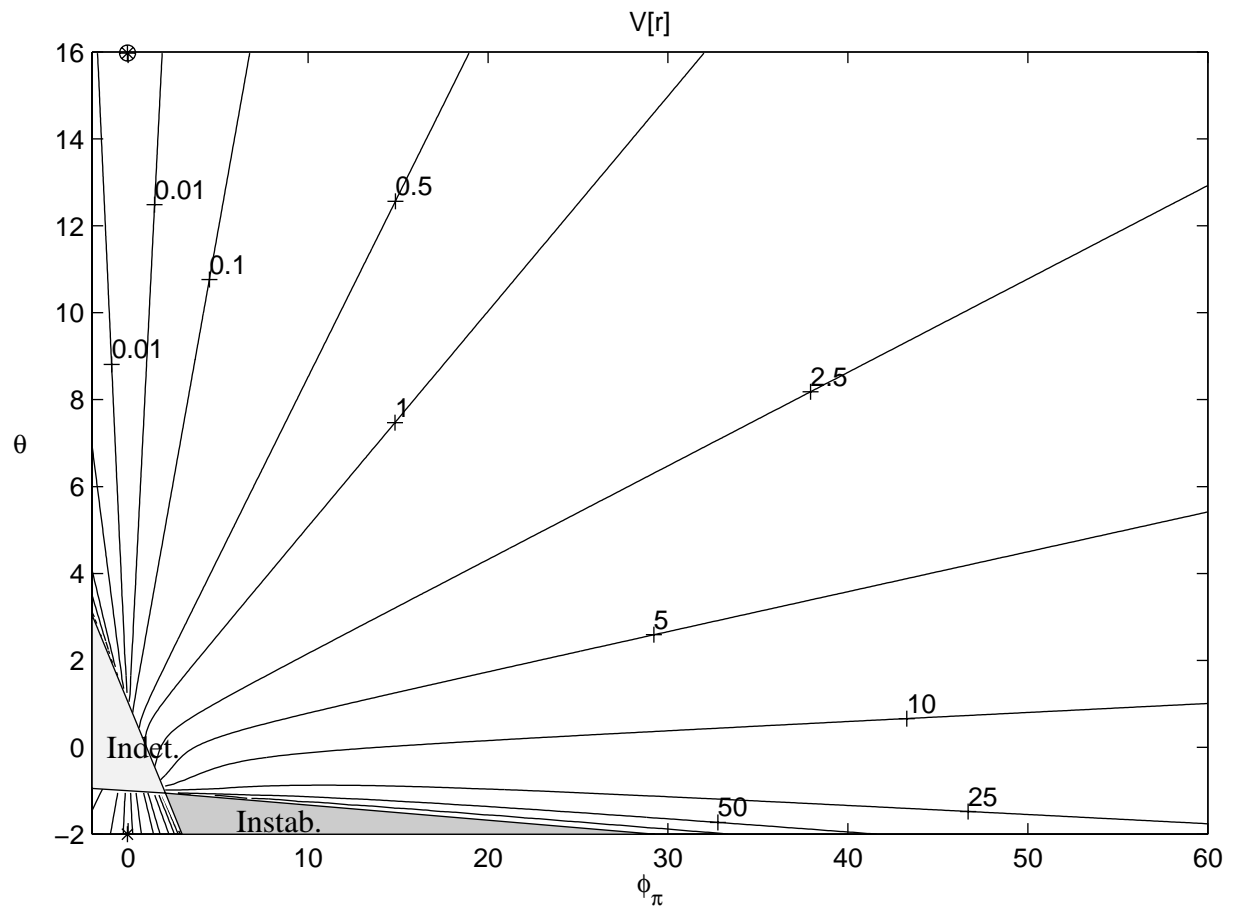


Figure 15

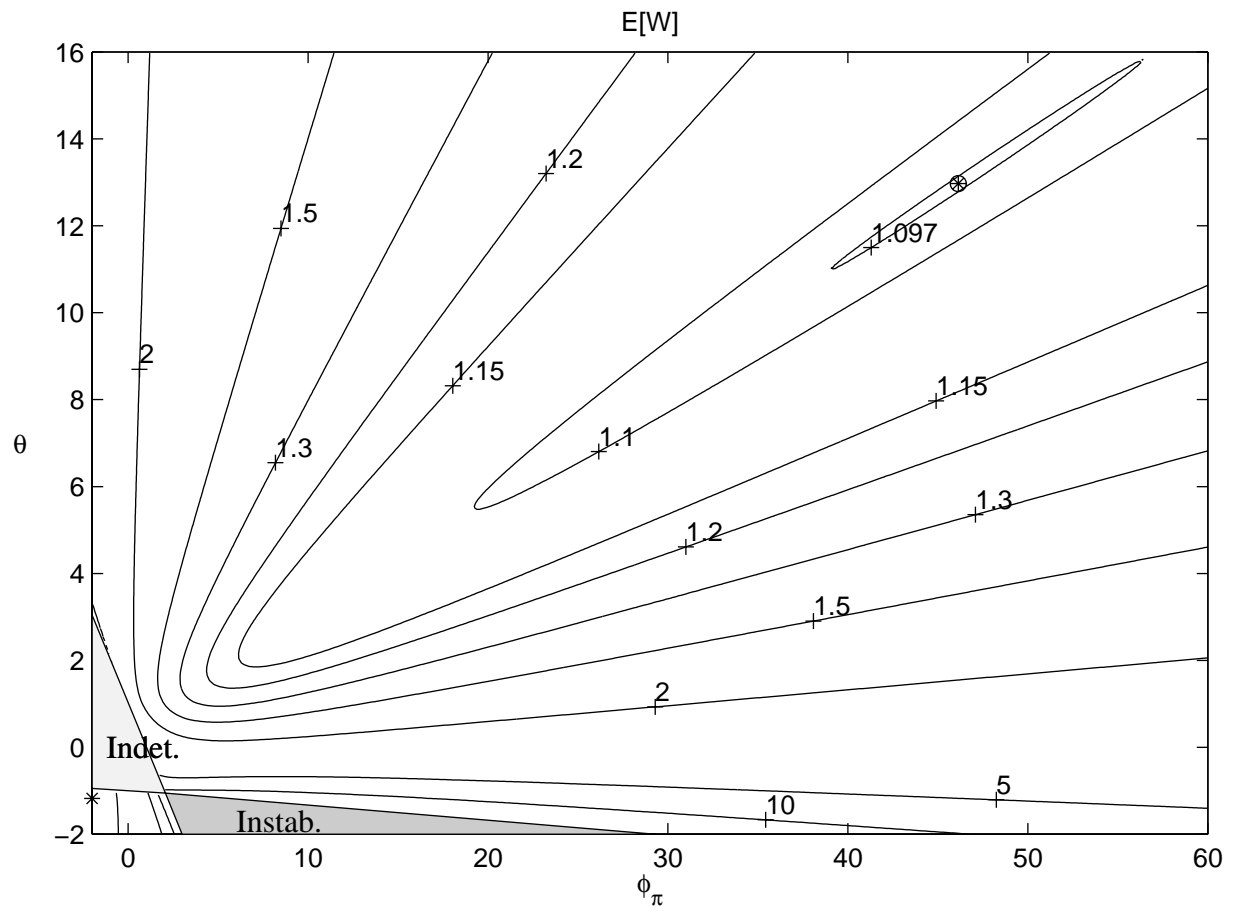


Figure 16