Timeless Perspective vs. Discretionary Monetary Policy in Forward-Looking Models

Bennett T. McCallum and Edward Nelson

Recent analyses by Clarida, Gali, and Gertler (1999), Jensen (2002), Svensson and Woodford (2004), Walsh (2003), and especially Woodford (1999, 2000, 2003) have been highly productive in advancing the profession’s understanding of monetary policy that is intended to be optimal. Specifically, these papers emphasize the importance, for policy purposes, of the distinction between macroeconomic models (of private behavior) that are “forward looking”—i.e., have equations that include expectations of future values of endogenous variables—and those that are not. This distinction—applied to the structural form of the model—is of great theoretical significance, since models derived from optimizing analysis almost invariably include expectations of future variables. A major point of the cited literature is that there is, in forward-looking models, an inefficiency that results from discretionary policymaking, relative to that of a well-designed policy rule, that obtains in addition to the familiar inflationary bias. (The inflationary bias has been extensively discussed in a huge literature that typically uses non-forward-looking models.) This “dynamic loss” arising from discretionary monetary policy, which is implicit in earlier work by Currie and Levine (1993) among others, has been valuably emphasized in the cited papers, especially Woodford (1999).

There are many associated issues, nevertheless, that remain to be considered. One of these is the quantitative extent to which a policy rule of the type in question provides improved outcomes relative to (optimal) discretionary behavior. That magnitude depends, of course, on the model employed—its parameter values and general aspects of the specification. An exploration of these features is clearly warranted. A second and related topic, moreover, concerns the distinction proposed by Svensson (1997, 1999) between “targeting rules” and “instrument rules.” Is there, in fact, a major difference? Or can targeting-rule outcomes be closely approximated by instrument-rule procedures? Third, in the context of optimal policy-rule analysis, issues concerning operationality—stressed by McCallum and Nelson (1999)—arise naturally. Is the superiority of rule-based over discretionary policymaking enhanced or diminished by realistic specification of information available to the policymaker? Finally, how important is this newly emphasized dynamic loss compared with that from the more familiar inflationary bias that arises from discretionary policymaking?

In exploring these issues, we begin with an exposition of the basic analysis that emphasizes Woodford’s concept of a “timeless perspective” monetary policy and its relationship to previous concepts of rule-based policymaking.

BASIC ANALYSIS

As an illustrative framework, we begin with the forward-looking macroeconomic model that is used by Woodford (1999, 2000) and also is a special case of the models in Clarida, Gali, and Gertler (CGG) (1999) and Jensen (2002).1 This simplest version features only a forward-looking price adjustment or aggregate supply relation of the Calvo type, augmented with shocks that keep the current natural-rate level of output from being economically efficient.2

Below, in addition, we consider a more general

---

1 The latter two papers permit first-order autoregressive processes for the shock variables, which make their systems somewhat richer than that considered by Woodford, and also consider model variants that include lagged inflation and output-gap terms.

relation that, although less clean theoretically, yields implications that some analysts (e.g., Fuhrer, 1997) consider to be more realistic empirically. Denoting inflation (relative to its steady-state value) in period $t$ by $\pi_t$, and the output gap by $y_t$, the basic relation is the New Keynesian Phillips curve (NKPC):

$$\pi_t = \alpha y_t + \beta E_t \pi_{t+1} + u_t,$$

where $\alpha > 0$, $0 < \beta < 1$, and $u_t$ is the shock term. For simplicity, we initially assume that the process generating $u_t$ is white noise. The model that we (and the cited authors) have in mind actually also includes an optimizing IS-type demand relationship of the form

$$y_t = E_t y_{t+1} + b_1 (R_t - E_t \pi_{t+1}) + v_t, \quad b_1 < 0$$

where $R_t$ is the central bank’s interest rate instrument and $v_t$ is a preference or government spending shock. But we shall at first pretend that the central bank can directly control $\pi_t$ as an instrument—an assumption that is very common in the literature and is innocuous in the present context. We will later extend the analysis in a manner that involves inclusion of equation (2) and use of an interest rate instrument.

The central bank’s objective function at time $t$ is taken to be of the form

(3) \[ \text{Minimize } E_t \sum_{j=1}^{\infty} \beta^j (\pi_{t+j}^2 + \alpha y_{t+j}^2), \quad \alpha > 0 \]

which Woodford (2003) has shown to be consistent with individual optimality in terms of agents’ preferences under certain reasonable conditions. Consequently, the central bank’s problem at some point in time, here taken (without loss of generality) to be $t = 1$, can be expressed as minimization of the Lagrangian expression

$$L_t = E_t [ (\pi_t^2 + \omega y_t^2) + \beta (\pi_{t+1}^2 + \omega y_{t+1}^2) + ... + \lambda_1 (\alpha y_t + \beta \pi_t + u_t - \pi_t) + \beta \lambda_2 (\alpha y_{t+1} + \beta \pi_{t+1} + u_{t+1} - \pi_{t+1}) + ... ]$$

with respect to $\pi_1, \pi_2, ...$, and $y_1, y_2, ...$ As shown by Woodford (1999) and CGG (1999), under policy commitment the optimizing conditions include

(5a) $2 \alpha y_t + \alpha \lambda_t = 0, \quad t = 1, 2, ...$

(5b) $2 \pi_t + \lambda_{t-1} - \lambda_t = 0, \quad t = 2, 3, ...$

(5c) $2 \pi_1 - \lambda_1 = 0.$

Here equations (1), (5a), (5b), and (5c) apparently determine optimal values of $\pi_t, y_t$, and $\lambda_t$ for period $t = 1$ and planned values as of $t = 1$ for periods $t = 2, 3, ...$. But these choices entail dynamic inconsistency, since the central bank could re-solve the problem in period 2 and would then choose $2 \pi_2 - \lambda_2 = 0$ instead of the condition $2 \pi_2 + \lambda_2 - \lambda_2 = 0$ that is suggested by (5b). Thus the standard commitment solution, in which the central bank implements (5a), (5b), and (5c), views the central bank as selecting values in $t = 2, 3, ...$ that it currently considers undesirable from the perspective of its own decision-making process. Since such a pattern of behavior seems highly implausible, this type of commitment solution does not provide an attractive equilibrium concept.

There is another equilibrium concept, however, involving a different type of commitment, that is much more attractive—as Woodford (1999) argues convincingly. Instead of using (1), (5a), and (5b) together with the start-up condition (5c) to determine paths of $\pi_t, y_t$, and $\lambda_t$ for $t = 1, 2, ...$, the central bank can use (1), (5a), and (5b) without any start-up condition by applying (5b) in all periods. This approach, which Woodford terms the “timeless perspective,” involves ignoring the conditions that prevail at the regime’s inception—say, by imagining that the decision to apply (5a) and (5b) had been made in the distant past. In this case, there is no dynamic inconsistency in terms of the central bank’s own decisionmaking process. Specifically, if there is no

---

5 The coefficient $\beta$ in (1) represents the private sector’s discount factor. Since its value is smaller than 1.0, equation (1) would imply that the Phillips curve is not of the accelerationist type if $\pi_t$ were defined as inflation. With our interpretation, which Woodford (1999) and CGG (1999) suggest, the central bank’s decision to apply (5a) and (5b) had been made in the distant past. In this case, there is no dynamic inconsistency in terms of the central bank’s own decisionmaking process. Specifically, if there is no

6 The model that Woodford uses to derive this welfare function has no $u_t$ disturbance in the Phillips curve (1). Giannoni (2000) provides a rationalization for the $u_t$ term that would continue to imply the objective (3).

7 Note that we are using the same discount factor in both (1) and (3). Below we briefly mention an implication of differing values.

8 In (4), the terms $E_t \pi_{t+j}$ from (1) can be written without $E_t$ operators because $E_t E_t \pi_{t+j} = E_t \pi_{t+j}$, by the law of iterated expectations.

9 When period $t+j$ comes around, the central bank can by assumption observe $y_{t+j}$ and $\pi_{t+j}$, so it can implement (5a)-(5b) exactly.
change in the central bank’s model, then the relationships between \( \pi_2 \) and \( Y_2 \) chosen by this process in period 2 agree with the relationship planned in period 1.

An alternative description of this mode of policy behavior can be obtained by specifying that the analyst’s concern is with macroeconomic performance within and across regimes, not with transitions from one regime to another. In this case, the analysis specifies that the policy regime has been in effect long enough that effects of the particular initial conditions, which obtained at the time of its inception, have become negligible. This is the conception adopted by Lucas (1980, p. 205), Taylor (1979, p. 1278), and others. Our contention is that this is the most appropriate presumption for monetary policy analysis. To us it seems implausible that, following a policy regime change, private agents could immediately begin forming expectations consistent with the new regime. The basic rational expectations approach requires that a policy regime has been in effect long enough for private agents to understand it and believe in its continuation.

It is perhaps worth mentioning that this timeless-perspective type of policy behavior agrees in spirit with what has been viewed by most analysts, since publication of Barro and Gordon’s (1983) exposition of the Kydland and Prescott (1977) insights, as “policymaking according to a rule.” The various quotes in Woodford (1999) taken from McCallum (1999a) illustrate that agreement, as does Woodford’s placement of his analysis in a section of his (1999) paper entitled “Rule-Based Policy-Making.” The modification that King and Wolman (1999, pp. 374-75) make to the commitment case, in their study of optimal monetary policy, also corresponds to adoption of a timeless perspective. It is also worth emphasizing that many studies of optimal monetary policy in forward-looking models have considered policies that are labeled “commitment,” but which (because these policies ignore the period 1 first-order condition and use only the remaining portion of the commitment conditions) should really be regarded as reflecting timeless perspective policy. Recent examples in this last category of studies include CGG (1999) and Batini and Nelson (2001).

For comparison, we need to derive the counterpart of conditions (5a)-(5c) provided by “discretionary” policymaking, i.e., a process that presumes period-by-period reoptimization involving each period’s start-up conditions. In this case the derivatives with respect to the terms in the Lagrangian expression (4) that correspond to \( E_t \pi_{t+1} \) in (1) are all equal to zero. Thus the counterpart of (5b) becomes

\[
2 \pi_t - \lambda_t = 0, \quad t = 1, 2, \ldots
\]

which is similar to the first-period condition (5c) in the commitment optimization, but now applies to each period. Note that discretion can be characterized by the absence of the lagged Lagrange multiplier in the central bank’s first-order condition, as stressed by Woodford (2003, Chap. 7).

In addition, let us express the policy-optimality conditions with the Lagrange multipliers \( \lambda_t \) substituted out. Then, for the discretionary optimum, we obtain from (6) and (5a) the following:

\[
\pi_t = -(\omega/\alpha) Y_t.
\]

By contrast, the timeless-perspective, rule-based condition implied by (5b) and (5a) is

\[
\pi_t = -(\omega/\alpha) (Y_t - Y_{t-1}).
\]

The latter expression is equivalent to (8) or (7) in Woodford (1999) and to (4.18) of CGG (1999). It is of some interest to note that in the special case \( \omega = \alpha \), and with constant potential output growth, the timeless-perspective rule (8) calls for nominal income growth targeting. This point is related to the findings reported by Jensen (2002) and Walsh (2003).

Quite recently, it has been recognized that use of (8) in all periods, as proposed by CGG (1999) and Woodford (1999), is not optimal within the class of time-invariant policy rules. Specifically, there is a slightly different rule that generates superior results on average, i.e., that yields a smaller unconditional expectation of the conditional expectation in (3). Analysis conducted to date suggests, however, that the welfare improvements are not substantially great.

---

10 See, for example, Woodford’s (1999) footnote 22.

11 King and Wolman’s (1999) modification is patterned after an analogous procedure in Rydland and Prescott’s (1980) study of optimal tax policy.

12 The reason is somewhat more complex than in the Barro-Gordon (1983) model, which is not forward looking; see Woodford (1999, pp. 308-9) or CGG (1999, p. 1672).

13 See Jensen and McCallum (2002), Jensen (2001), and Blake (2001). These papers indicate that optimality requires that (8) be altered to \( \pi_t = -(\omega/\alpha) (\gamma_t - \beta_\gamma Y_{t-1}) \). If private-sector and central-bank discount factors differ, it is the private-sector value that appears in this latter expression.
In any event, these results do not negate interest in comparisons between the Woodford-CGG timeless-perspective results and those based on discretionary behavior.

**Equilibrium Behavior in the Basic Model**

To determine how inflation and the output gap behave in the timeless-perspective equilibrium, we obtain the rational expectations solution to the model consisting of the policy rule (8) and the private behavioral relation (1). In particular, we look for the minimal state variable (MSV) solution that excludes bubbles and sunspots, as discussed by McCallum (1999b). Thus we conjecture that \( \pi_t \) and \( y_t \) are related to the clearly relevant state variables \( y_{t-1} \) and \( u_t \) as follows:

(9a) \[ \pi_t = \phi_{11} y_{t-1} + \phi_{12} u_t, \]

(9b) \[ y_t = \phi_{21} y_{t-1} + \phi_{22} u_t. \]

Then \( E_t \pi_{t+1} = \phi_{11}(\phi_{21} y_{t-1} + \phi_{22} u_t) \) and substitution into (1) and (8) yields the undetermined-coefficient relationships:

(10a) \[ \phi_{11} = \alpha \phi_{21} + \beta \phi_{12}, \]

(10b) \[ \phi_{12} = \alpha \phi_{22} + \beta \phi_{12} + 1, \]

(10c) \[ \phi_{11} = (\omega/\alpha)(1 - \phi_{21}), \]

(10d) \[ \phi_{21} = -\omega \phi_{22}. \]

From (10a) and (10c), we find that \( \phi_{21} \) satisfies

(11) \[ \beta \phi_{21}^2 - \gamma \phi_{21} + 1 = 0, \]

where \( \gamma = (1 + \beta + \omega^2/\alpha). \) The relevant root, according to both the stability and MSV criteria, is

(12) \[ \phi_{21} = [\gamma - (\gamma^2 - 4\beta)^{0.5}] / 2\beta, \]

which satisfies \( 0 < \phi_{21} < 1. \) Following CGG (1999), we use the symbol \( \delta = \phi_{21}. \) Then the values for \( \phi_{11}, \phi_{12}, \) and \( \phi_{22} \) can be found to be \( \phi_{11} = (\omega/\alpha)(1 - \delta), \phi_{12} = 1/(\gamma - \beta \delta), \) and \( \phi_{22} = -\omega \alpha \phi_{22} \gamma / \beta \delta, \) and the solutions are

(13) \[ \pi_t = (\omega/\alpha)(1 - \delta)y_{t-1} + (\gamma - \beta \delta)^{-1}u_t, \]

and

(14) \[ y_t = \delta y_{t-1} - [(\omega/\alpha)(\gamma - \beta \delta)]u_t. \]

These can be shown, with some tedious algebra, to agree with solution expressions reported by CGG (1999, e.g., their equation (8.1)).

Finally, to find the MSV equilibrium under discretionary optimal policy, we use (7) rather than (8) as the policy rule. In a system consisting of (1) and (7), there are no clearly relevant state variables other than \( u_t, \) so we conjecture a solution of the form

(15) \[ \pi_t = \phi_1 u_t, \]

(16) \[ y_t = \phi_2 u_t. \]

Then \( E_t \pi_{t+1} = 0 \) and the values of \( \phi_1 \) and \( \phi_2 \) are found to be \( \omega \alpha(\omega + \alpha^2) \) and \( -\alpha \omega(\omega + \alpha^2), \) respectively.

Neither Woodford (1999) nor CGG (1999) includes an analysis of the relative losses—the unconditional expectations of the objective function—under the two modes of policymaking. Indeed, they do not actually put forth any claim that the timeless-perspective losses are generally smaller than those from discretionary policymaking. We do not attempt any general algebraic analysis here, but proceed by examining the issue quantitatively using calibrated models with specific parameter values varied over fairly wide but realistic intervals. Such an analysis is included in the next section.

**QUANTITATIVE ANALYSIS**

Our agenda now is to specify values for the model’s parameters \( \alpha, \beta, \) and \( \omega, \) find the rational expectations solutions described above; and report, for a range of values for the variances and serial correlation of \( u_t, \) the average values of the loss function. The average values of the intertemporal loss function (3) are proportional to the mean of the instantaneous loss function—its unconditional expectation—which is what we report. (Thus the

---

14 Issues involving determinacy of this solution, and others considered below, are considered by Bullard and Mira (2003) and Svensson and Woodford (2004). More fundamental, in our opinion, is the learnability of various solutions, also discussed by Bullard and Mira. In that regard, Evans and Honkaphoja (2003) have shown that the solution (13)-(14) is learnable if an appropriate interest rate rule is used to implement the optimality condition (8), but is not learnable for some other modes of implementation. Similar results apply to other solutions discussed below.

15 The recent results of Blake (2001) indicate that such a claim would be incorrect, although the contrary cases involve unusually low values of \( \alpha \omega \) and \( \beta. \)

16 A few quantitative results have previously been reported by Giannoni (2000), Vestin (2000), Woodford (1999), and Walsh (2003), but without the type of systematic exploration provided here.

17 Here we follow the example of King and Wolman (1999), Rotemberg and Woodford (1999), Rudebusch and Svensson (1999), and Walsh (2003) in our use of the unconditional expectation of (3) as the policy criterion.
unconditional expectation of (3) equals the reported values multiplied by $1 + \beta + \beta^2 + \ldots = (1 - \beta)^{-1}$. In what follows, these values are calculated using asymptotic formulae for the moments of the variables in the model (e.g., Hamilton, 1994, p. 265).\footnote{These values have been checked by comparison with averages of the same statistics across 100 stochastic simulations (200 periods).} We use our modification of the QZ algorithm of Klein (2000) to obtain the MSV solution.

**Results for the Basic Model**

Table 1 reports values of the loss function for a range of $\alpha$ and $\omega$ values, with $\beta$ kept at 0.99 throughout. For $\alpha$, we suggest that actual values probably lie between 0.01 and 0.05; see, e.g., the estimates in Galí and Gertler (1999). For the central bank preference parameter $\omega$, our range of 0.001 to 0.1 includes values that place almost all weight on inflation variability and values that give much weight to output gap variability. Since we are using quarter-year time periods, equal weights in terms of annualized inflation (as in the original Taylor rule) imply $\omega = (1/4)^2 = 0.0625$. The standard deviation of the white noise $u_t$ shocks is taken to be 0.005 (i.e., 0.5 percent).\footnote{The value chosen for this standard deviation directly influences the values of calculated losses, but does not influence the relative magnitudes of the losses under TP and discretionary policies.} Thus the annualized standard deviation is about 2 percent, slightly less than is realistic for the U.S. economy. In each entry of Table 1 there are two numbers; the first is the average (i.e., unconditional expectation) loss for the timeless-perspective (TP) solution, and the second is for the discretionary (DIS) solution. From the table it can be seen that the TP policy produces smaller losses than the DIS policy for all examined values of $\alpha$ and $\omega$. The quantitative extent of the difference is about 15 to 20 percent for most values in the table, but falls to a magnitude as low as 2 percent.

What, if anything, can be said about the absolute levels of these loss magnitudes? Clearly it is possible to compare them to the costs of a steady, maintained inflation. Suppose that inflation is kept constantly at a rate that exceeds the target rate by 1 percent per year. That implies $\pi_t = 0.01/4 = 0.0025$, so with $y_t$ kept at zero, the value of the period loss becomes $0.0025^2 = 0.63 \times 10^{-5}$. Thus the TP and DIS losses in Table 1 with $\alpha = 0.05$ and $\omega = 0.01$, for example, are about 2.4 and 3.2 times as large as the cost of an inflation rate that exceeds the target by 1.0 percent per year. Translating the inflation cost into equivalent consumption terms would be highly problematic, since the results depend sensitively upon the precise shape of the money demand function in the vicinity of the target inflation rate. This problem would remain even if we were to assume that the central bank loss function were based on the utility function of individual agents.

**Results for the Basic Model with Serially Correlated Shocks**

To consider whether these results are robust, we modify the model somewhat. In particular, we now assume that the $u_t$ shock process is serially correlated according to a first-order autoregressive specification with an autoregression parameter value of 0.8. This change does not affect the TP and DIS rules for the basic model, but will result in solution processes for inflation and the output gap that feature considerable persistence, much more like actual data than those generated by the basic model with white noise $u_t$ shocks. We retain a value of 0.005 for the standard deviation of $u_t$ by reducing the innovation variance by a factor of $[1/(1 - 0.8^2)] = 2.778$. Results are shown in Table 2.

A greater percentage difference now holds in the TP and DIS outcomes for most $\alpha$ and $\omega$ values. The ratios of DIS to TP losses, that is, are somewhat

---

<table>
<thead>
<tr>
<th>Value of $\alpha$</th>
<th>Value of $\omega$</th>
<th>Value of $\pi_t$</th>
<th>Value of $\omega_t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.001</td>
<td>0.01</td>
<td>0.0625</td>
</tr>
<tr>
<td></td>
<td>0.21/0.23</td>
<td>0.96/1.25</td>
<td>1.69/2.16</td>
</tr>
<tr>
<td>0.05</td>
<td>0.59/0.71</td>
<td>1.54/2.00</td>
<td>2.07/2.40</td>
</tr>
<tr>
<td>0.01</td>
<td>1.84/2.27</td>
<td>2.28/2.48</td>
<td>2.43/2.50</td>
</tr>
</tbody>
</table>

Table 1

Losses with TP and DIS Policy Behavior, Basic NKPC
(Reported values are losses times $10^5$, TP/DIS)
larger than in the case with white noise shocks. This is not too surprising, for the fundamental advantage of the TP rule is that it takes correct account of private sector expectations and, therefore, of intertemporal aspects of the situation, which are more pronounced when serial correlation of the shocks is included.

**Alternative Model**

An alternative specification that also tends to generate persistence in inflation, and consequently has been prominent in recent research, is provided by replacement of price-adjustment relation (1) with the following:

\[(17) \pi_t = \alpha y_t + \beta \theta \pi_{t+1} + \beta (1-\theta)\pi_{t-1} + u_t. \quad 0 < \theta < 1\]

Relations of this general type have been promoted by Fuhrer (1997), among others, and are considered in the rule analysis of CGG (1999), Jensen (2002), and Walsh (2003). To find the TP policy rule with (17) replacing (1), we follow the procedure outlined earlier and obtain the following first-order conditions in place of (5):20

\[(18a) 2\omega y_t + \alpha \lambda_t = 0 \quad t = 1, 2, \ldots\]
\[(18b) 2\pi_t + \theta \lambda_{t-1} - \lambda_t + \beta^2 (1-\theta)\pi_{t+1} = 0 \quad t = 2, 3, \ldots\]
\[(18c) 2\pi_t - \lambda_t + \beta^2 (1-\theta)\lambda_2 = 0.\]

Adopting the Woodford-CGG timeless perspective approach, by substituting out the \(\lambda_t\) multipliers between (18a) and (18b), yields the optimality condition

\[(19) \pi_t = (\omega/\alpha)[\theta y_{t-1} - y_t + \beta^2 (1-\theta)\pi_{t+1}]. \quad t = 1, 2, \ldots\]

Here, \(E_\pi y_{t+1}\) appears instead of \(y_{t+1}\), because the latter is not known at \(t\).

For the case of discretionary optimization, there are actually two possible concepts. First, one might conceive of the central bank as implementing (18a) and (18c) in period 1 and planning to implement (18a) and (18b) in each subsequent period. When period 2 arrives, however, the central bank re-solves its problem and again implements (18a) and (18c), now updated to period 2. Indeed, in this case the central bank re-solves and implements this solution in each period. With Lagrange multipliers substituted out, the relevant optimality condition is

\[(20) \pi_t = -\frac{\omega}{\alpha}[y_t - \beta^2 (1-\theta)E_\pi y_{t+1}].\]

where again it is recognized that \(y_{t+1}\) is not known in period \(t\). The second concept, used by CGG (1999, p. 1692) and Jensen (2002), does not involve the dynamic inconsistency that is clearly implied by the first. Instead of planning to implement (18b) in future periods, the central bank recognizes in period 1 that in period 2 it will behave just as it does in period 1. Therefore, in minimizing (3), \(E_\pi \pi_2\) in the constraint (17) for period 1 will be replaced with \(\rho_1 \pi_1\), where \(\rho_1\) is a parameter of the equilibrium solution expression \(\pi_t = \rho_1 \pi_{t-1} + \rho_1 u_t\). In the present case with white noise \(u_t\), the relevant optimality condition with this conception of discretionary behavior is

\[(21) \pi_t = -\frac{\omega}{\alpha}[(1-\beta \theta \rho_1) y_t - \beta^2 (1-\theta)E_\pi y_{t+1}].\]

Thus there is a smaller responsiveness of inflation (and a larger responsiveness of output) to shocks than would be present if policy behavior were as implied by (20). Since (21) evidently reflects a more

---

20 In general, changing the Phillips curve specification means that the loss function (3) can no longer be obtained directly from an approximation of household utility. For example, Steinsson (2003) shows that a Phillips curve like equation (17) implies that the period loss function is no longer time-separable. Following CGG and Jensen, we neglect this nonseparability and continue to use (3) as our welfare criterion.
standard version of discretion than (20), it will be used in what follows.\(^{21}\)

Results are reported in Table 3 for the case in which (17) describes price adjustment behavior, with \(\theta = 0.5\), when \(u_t\) is white noise. Here the ratio of DIS losses to TP losses is somewhat larger than in Table 1 for the lower right-hand cells but smaller elsewhere. In all cases covered by our \(\alpha\) and \(\omega\) values, the TP losses are smaller than the DIS losses. The table therefore establishes, for a wide range of parameter choices, that the superiority of TP over DIS policies is robust to allowing for intrinsic persistence in inflation.

### TARGET RULES AND INSTRUMENT RULES

A monetary policy that implemented the optimality conditions studied above would correspond to following what Svensson (1997, 1999) terms “targeting rules,” as distinct from “instrument rules.” In these papers, as well as others, Svensson argues that consideration of targeting rules is preferable for actual central banks and for policy analysis. McCallum (1999a) and McCallum and Nelson (1999) have, by contrast, suggested that instrument rules are more interesting from a normative point of view. It could also be argued that they are more relevant empirically, in the sense that the actual inflation-targeting regimes currently in place in New Zealand, Canada, the United Kingdom, and elsewhere are more satisfactorily represented by formal analytical models with instrument rules than with target rules.\(^{22}\)

An important part of this argument is that no actual central bank has revealed what its loss function is—e.g., what its value of \(\omega\) is in expression (3). Of course an argument of this nature can never be conclusive, but we would point out that Woodford (1999, pp. 287-99) has presented a sophisticated discussion that predominantly supports this position.

A strictly analytical claim made by McCallum (1999a, p. 1493, footnote 17) is that an instrument rule can typically be written so as to imply instrument responses that would tend to bring about the satisfaction of any (feasible) specified target rule. In the context of the present analysis, for example, one could include the optimizing IS relation (2) as part of the model and then specify an instrument rule for \(R_t\) that is designed to implement an optimality condition such as (8). In this case, the rule could be written as

\[(22)\]

\[R_t = (1-\mu_2)\{\bar{r} + \pi_t + \mu_1[\pi_t + (\omega/\alpha)(y_t - y_{t-1})]\} + \mu_2 R_{t-1},\]

which, with \(\mu_1 > 0\), \(\mu_2 \geq 0\), is similar to an extended\(^{23}\) version of the Taylor (1993) rule, but with \(\pi_t + (\omega/\alpha)(y_t - y_{t-1})\) rather than \(\pi_t + y_t\) as the target variable, i.e., the variable that the rule seeks to keep close to some desired value. If the economy is one in which current aggregate demand can be influenced by \(R_t\), then as \(\mu_1\) is increased, the variability of the term in square brackets in (21) tends to be decreased, yielding an approximation to the satisfaction of optimality condition (8).\(^{24}\)

To determine whether it is in fact the case that increasing \(\mu_1\) values would lead to approximate satisfaction of (8)—and likewise of the discretionary

\[\text{Table 3}\]

**Losses with TP and DIS Policy Behavior, Model Including (17)**

*(Reported values are losses times 10\(^5\), TP/DIS)*

<table>
<thead>
<tr>
<th>Value of (\alpha)</th>
<th>0.001</th>
<th>0.01</th>
<th>0.0625</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.22/0.23</td>
<td>1.36/1.48</td>
<td>3.70/4.34</td>
<td>4.56/5.45</td>
</tr>
<tr>
<td>0.05</td>
<td>0.72/0.75</td>
<td>2.99/3.44</td>
<td>6.60/8.17</td>
<td>7.83/9.85</td>
</tr>
<tr>
<td>0.01</td>
<td>4.56/5.45</td>
<td>10.6/13.7</td>
<td>17.7/22.6</td>
<td>19.7/24.6</td>
</tr>
</tbody>
</table>

---

\(^{21}\) We proceed computationally by assuming a value for \(\rho_1\), solving the model conditional on that value, determining the value implied by the solution, and iterating. For an alternative, dynamic programming approach to the problem, see Steinsson (2003).

\(^{22}\) See, for example, the discussions of the respective central bank practices given by Archer (2000), Freedman (2000), and King (1999).

\(^{23}\) The extension takes the form of an added \(R_{t-1}\) term to reflect interest rate smoothing.

\(^{24}\) This argument does not maintain that (22) is the only instrument rule that would serve the purpose of implementing (8), but merely that it will do so (and has been mentioned in the literature).
optimality condition (7)—consider the values reported in Table 4. There, $\alpha = 0.05$ and $\mu_2 = 0$ are retained throughout, with various values of $\omega$ specified and $\mu_1$ increased from the Taylor value 0.5 to extremely large magnitudes. The shock term in relation (2) includes two components, a white noise taste component with standard deviation 0.02 and also $y_t - E_t \{y_{t+1}\}$, where the natural-rate value $y_t$ comes from an AR(1) process with AR parameter 0.95 and innovation standard deviation 0.007. The results indicate that, at least for this example, the instrument rule approximates very closely the target-rule optimality conditions for large $\mu_1$ values (strong feedback responses). With $\mu_1 \geq 50$, for example, the $\omega = 0.0625$ case gives TP and DIS loss values of 2.09 and 2.41, respectively, which are essentially identical to the target-rule losses shown in Table 1. Thus, instrument rules can be written to include target rules as extreme special cases, but are more general.

### OPERATIONALITY

Exercises such as those of the preceding sections are interesting and even enlightening, but are far removed from the monetary policy problems facing actual central bankers. In reality, decisionmakers at central banks have only vague notions about the "true model"—i.e., the workings of the actual economy—and have highly incomplete and imperfect information regarding current values of many variables of macroeconomic importance. Serious studies of desirable policy should recognize these features of reality. We now wish to determine how such operationality considerations are related to the issues regarding optimality in forward-looking models that have been considered here. Clearly, a complete study is beyond the scope of this paper, but some leading problems can be considered. First we consider two particular points, ones that have been stressed in previous work by McCallum and Nelson (1999) and McCallum (1999a).

The first point is the central bank’s lack of knowledge of the current value of output at the time it sets its interest rate instrument for that period. To be more realistic, one could include the most recent period’s value $y_{t-1}$, but a preferable approach would be to use $E_{t-1}y_t$. Accordingly, we now investigate the effects of including $E_{t-1}y_t$ in place of $y_t$ in instrument rule simulations such as those used earlier. In addition, we consider cases in which current inflation is not observed, so that $E_{t-1}\pi_t$ is used by the central bank in place of $\pi_t$, and in which neither of these variables is observed.

A first set of results is shown in Table 5, where the first row repeats results from Table 4 for comparison. The second row gives the results with the expected current output gap included in place of the (unobserved) current value. It will be seen that the magnitude of the losses seen in the table is, in this case, much greater than with full information, with the extent of the increase positively related to $\omega$ (i.e., to the strength of the response to the imperfectly observed gap variable). For each $\omega$ value con-

---

### Table 4

<table>
<thead>
<tr>
<th>Value of $\omega$</th>
<th>0.001</th>
<th>0.01</th>
<th>0.0625</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value of $\mu_1$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>2.51/2.51</td>
<td>2.90/2.86</td>
<td>5.29/4.20</td>
<td>7.36/4.69</td>
</tr>
<tr>
<td>5.0</td>
<td>2.12/2.14</td>
<td>2.45/2.43</td>
<td>3.38/2.63</td>
<td>3.40/2.62</td>
</tr>
<tr>
<td>50.0</td>
<td>0.86/1.04</td>
<td>1.58/2.03</td>
<td>2.09/2.41</td>
<td>2.17/2.44</td>
</tr>
<tr>
<td>500.0</td>
<td>0.59/0.72</td>
<td>1.54/2.00</td>
<td>2.07/2.40</td>
<td>2.16/2.44</td>
</tr>
</tbody>
</table>
Table 5

**TP and DIS Losses with Unobservable Output**

Basic Model with $\alpha = 0.05$, $\mu_2 = 50$, and $\mu_2 = 0$

(Reported values are losses times $10^5$, TP/DIS)

<table>
<thead>
<tr>
<th>Value of $\omega$</th>
<th>0.001</th>
<th>0.01</th>
<th>0.0625</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>With $y_t$ and $\pi_t$ in rule</td>
<td>0.86/1.04</td>
<td>1.58/2.03</td>
<td>2.09/2.41</td>
<td>2.17/2.44</td>
</tr>
<tr>
<td>With $E_t y_t$ and $\pi_t$ in rule</td>
<td>0.74/0.90</td>
<td>3.21/3.29</td>
<td>12.0/17.2</td>
<td>16.3/27.1</td>
</tr>
<tr>
<td>With $y_t$ and $E_{t-1} \pi_t$ in rule</td>
<td>2.62/2.57</td>
<td>2.58/2.52</td>
<td>2.53/2.50</td>
<td>2.52/2.50</td>
</tr>
<tr>
<td>With $E_{t-1} y_t$ and $E_{t-1} \pi_t$ in rule</td>
<td>2.77/2.64</td>
<td>12.5/3.00</td>
<td>185.6/5.10</td>
<td>31,220/6.60</td>
</tr>
<tr>
<td>With $E_{t-1} y_{t+1}$, $E_{t-1} y_t$, and $E_{t-1} \pi_t$ in rule</td>
<td>2.64/2.64</td>
<td>3.00/3.00</td>
<td>5.10/5.10</td>
<td>6.60/6.60</td>
</tr>
</tbody>
</table>

Table 6

**TP and DIS Losses with Unobservable Output**

Model Including (17) with $\alpha = 0.05$, $\mu_1 = 50$, and $\mu_2 = 0$

(Reported values are losses times $10^5$, TP/DIS)

<table>
<thead>
<tr>
<th>Value of $\omega$</th>
<th>0.001</th>
<th>0.01</th>
<th>0.0625</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>With $E_{t-1} y_{t+1}$, $y_t$, and $\pi_t$ in rule</td>
<td>1.03/1.11</td>
<td>3.05/3.58</td>
<td>6.60/8.24</td>
<td>7.84/9.88</td>
</tr>
<tr>
<td>With $E_{t-1} y_{t+1}$, $E_{t-1} y_t$, and $\pi_t$ in rule</td>
<td>0.90/0.99</td>
<td>4.58/4.56</td>
<td>28.3/25.0</td>
<td>44.5/39.7</td>
</tr>
<tr>
<td>With $E_{t-1} y_{t+1}$, $y_t$, and $E_{t-1} \pi_t$ in rule</td>
<td>2.31/2.41</td>
<td>3.92/4.41</td>
<td>7.06/8.59</td>
<td>8.20/10.2</td>
</tr>
<tr>
<td>With $E_{t-1} y_{t+1}$, $E_{t-1} y_t$, and $E_{t-1} \pi_t$ in rule</td>
<td>2.22/2.35</td>
<td>4.39/4.58</td>
<td>25.3/11.0</td>
<td>51.2/14.2</td>
</tr>
<tr>
<td>With $E_{t-1} y_{t+1}$, $E_{t-1} y_{t+1}$, and $E_{t-1} \pi_t$ in rule</td>
<td>2.89/3.12</td>
<td>3.39/4.36</td>
<td>8.47/17.3</td>
<td>11.0/26.4</td>
</tr>
</tbody>
</table>

NOTE: $y_{t-1}$ appears in all the TP rules except those of the last row, where it is replaced by $E_{t-1} y_t$.

Considered, it remains true that the TP losses are smaller than the DIS losses. In the third row, we suppose that inflation (instead of output) is currently unobservable. In this case, the losses are essentially equal for all $\omega$ values and for both TP and DIS policies. The value of the loss function, moreover, is very nearly equal to the value of the variance of the $u_t$ shock term.

In the fourth row, we suppose that both inflation and output are currently unobservable. In this case, the TP losses jump up drastically while the DIS losses increase, but by much less. It is understandable that losses could be very large in this case, for the setup is one in which policy is in effect trying to stabilize current variables although they are not observable. In the discretionary case, certainty equivalence implies that the attempt at stabilization is being carried out as efficiently as possible when $t-1$ expectations are used in the rule in the absence of current observations (see Svensson and Woodford, 2003), but this principle does not carry over directly to the TP case. In that case, it turns out that inclusion of $E_{t-1} \pi_{t+l}$, $E_{t-1} y_{t+1}$, and $E_{t-1} y_t$ yields much better results. In fact, those results, shown in the fifth row of Table 5, are equivalent to those given in the fourth-row DIS cases.

A second set of results, pertaining to the case in which the price adjustment relation (17) replaces (1), is given in Table 6. Qualitatively, the results are not too different from those of Table 5. In particular, when neither $\pi_t$ nor $y_t$ is currently observable, the TP performance is poor. But it can be improved by shifting forward the dates of each variable (whose values are those expected on the basis of $t-1$ data).
Indeed, in this case, the TP results are superior to those based on the DIS procedure, instead of being equal as in Table 5.

Our second point of concern is arguably of even greater practical importance. It involves the unobservability of the natural-rate level of output that goes into the central bank’s measure of the output gap. In this case the nature of the problem is quite different, we contend. Rather than reflecting merely a lack of current information, the problem is largely conceptual—that is, it stems from the existence of various different concepts of the relevant reference value (which we have been calling “natural rate”). That there are several distinct concepts in use is implicit in the terms used by different researchers and practitioners. In addition to the term “potential,” which is frequently used by practitioners, there are the words “trend,” “capacity,” “NAIRU,” “market clearing,” and “flexible price,” besides “natural rate.” There are perhaps fewer distinct concepts than terms, but there seem to be at least three fundamentally different ones: trend, NAIRU, and flexible-price concepts. And, of course, there are many ways of measuring trend output that are quite different in their effects. Furthermore, since reliance on any particular concept will be maintained over time, differences will not possess the orthogonality properties of pure “noise.”

Which of the concepts is most appropriate theoretically? From the perspective of dynamic, optimizing analysis, the answer is the third of the three just listed, the flexible-price concept—that is, the output level that would prevail in the absence of nominal price stickiness. There have been very few attempts to implement this type of measure empirically, but there is one in McCallum and Nelson (1999), which we briefly review.

This procedure begins with the assumption that output is produced according to a Cobb-Douglas production function relating the log of output linearly to the logs of labor and capital \((n_t \text{ and } k_t)\), a deterministic trend, and a shock term \(a_t\), reflecting the stochastic component of technological change. Then, since \(k_t\) and \(a_t\) are given in \(t\) whether or not prices are flexible, the difference between the logs of actual and flexible-price output (i.e., the output gap) will be proportional to the difference between actual and flexible-price labor input, \(n_t - \bar{n}_t\). For simplicity, McCallum and Nelson (1999) assumed that the flexible-price level \(\bar{n}_t\) (per period, per person) is a constant and, numerically, they measured \(n_t\) for the United States, 1955:Q1–1996:Q4, as total man-hours employed in nonagricultural private industry divided by the civilian labor force. This measure is scaled so that the average value of \(n_t - \bar{n}_t\) equals zero. The necessity of that step is undesirable, but on the positive side there is no deterministic trend in the resulting \(n_t - \bar{n}_t\) series. Then, using 0.7 as the elasticity of output with respect to labor, they constructed a series for the output gap \(\bar{y}_t\) (shown in McCallum and Nelson, 1999, p. 28) and contrasted with a measure based on simple log-linear detrending. This series, in combination with the corresponding output series, provides a series for \(\bar{y}_t\). It has approximately the time series properties assumed above.

An important point is that non-zero realizations of the technology shock \(a_t\) affect the McCallum-Nelson measure of \(\bar{y}_t\) one-for-one, whereas many detrending procedures, used extensively by academics and to some extent by central banks, remove \(a_t\) almost entirely from each period’s measure of \(\bar{y}_t\). The same is true, furthermore, for many NAIRU-based procedures. So the question at hand is whether this conceptual discrepancy is of quantitative importance—whether the use of a mistaken concept would create major welfare losses from policy rules that rely upon measures of the output gap. We approach this question here by assuming that the McCallum and Nelson (1999) measure of the gap is correct, but that the central bank incorrectly uses the measure based on linear detrending in the context of instrument rule (22). For simplicity, we assume that the central bank has accurate knowledge of the true trend, which is excessively optimistic, so the conceptual error as implemented is only that the central bank neglects the influence of \(a_t\) on \(\bar{y}_t\).

Results are reported in Table 7. The loss values reported there can be compared with those in Table 4, in which the experiment is the same except for the postulated mis-measurement of \(\bar{y}_t\). It is clear that the consequences of the conceptual error are quite substantial, except for \(\omega = 0.001\), and are much larger for large values of \(\omega\). Because these values imply giving more weight to the output gap, the results are consistent with the suggestion of McCallum (1999a) and Orphanides (2003) that it is dangerous to respond strongly to measures of the output gap. Furthermore, Table 7 indicates that the TP outcomes are considerably more desirable

\[^{29}\] Gali and Gertler (1999) also use labor market data, in a different but related manner, in the context of estimating the Calvo specification (equation (1)).
than those resulting from DIS behavior. This result is in keeping with the spirit of the suggestions of McCallum (1999a), Orphanides (2003), and Jensen (2002) that responding to some variable reflecting nominal income growth may be more attractive than responding to the level of the output gap.

**INFLATIONARY BIAS**

An issue of obvious interest is how the magnitudes of the losses shown in Tables 1 through 7 compare with those implied by the discretionary inflationary bias discussed in the enormous literature that uses non-forward-looking models. The inflationary bias carries over to the forward-looking models, as Woodford (1999) and CGG (1999) have pointed out, if the central bank’s objective function includes terms such as \( \pi_t^2 + \omega(y_t - k)^2 \), with \( k > 0 \), reflecting a desire by the central bank to keep output above the natural-rate value that would obtain on average in the absence of nominal frictions (i.e., with fully flexible prices). In the model at hand, the magnitude of the bias is simply \( \omega(\alpha(k^2) \), as can be easily verified. To get a clear idea of the magnitudes involved, let us then suppose that \( k = 0.01 \), i.e., that the central bank aims for a level of output that exceeds the natural-rate value by 1 percent. Then if \( \omega \alpha = 1 \), the bias would be 0.01 and its square, 0.0001, would be appropriate for comparison with the values in Tables 1, 4, and 5.\(^30\) Those tables’ entries are losses multiplied by \( 10^5 \), of course, so in this case the loss value comparable to the first-row/fourth-column entries of Table 1 would be 10. More generally, we have the values reported in Table 8, where for values of \( \omega \alpha \) equal to or greater than 0.0625 the inflationary bias is more important, if relevant, than the newly emphasized dynamic loss.

---

**Table 7**

**Losses from Responding to Incorrect Concept for Potential Output**
(Reported values are losses times \( 10^5 \), TP/DIS)

<table>
<thead>
<tr>
<th>Value of ( \mu_1 )</th>
<th>Value of ( \omega )</th>
<th>0.001</th>
<th>0.01</th>
<th>0.0625</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>4.09/4.31</td>
<td>4.37/6.67</td>
<td>6.11/18.8</td>
<td>7.75/24.8</td>
<td></td>
</tr>
<tr>
<td>5.0</td>
<td>2.13/2.21</td>
<td>2.42/3.98</td>
<td>4.00/16.5</td>
<td>4.93/22.7</td>
<td></td>
</tr>
<tr>
<td>50.0</td>
<td>0.86/1.06</td>
<td>1.63/3.38</td>
<td>3.11/16.2</td>
<td>4.15/22.5</td>
<td></td>
</tr>
<tr>
<td>500.0</td>
<td>0.59/0.74</td>
<td>1.61/3.33</td>
<td>3.12/16.2</td>
<td>4.17/22.5</td>
<td></td>
</tr>
</tbody>
</table>

**Table 8**

**DIS Losses due to Inflation Bias**
Basic Model
(Reported values are losses times \( 10^5 \))

<table>
<thead>
<tr>
<th>Value of ( \alpha )</th>
<th>Value of ( \omega )</th>
<th>0.001</th>
<th>0.01</th>
<th>0.0625</th>
<th>0.10</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>0.001</td>
<td>0.10</td>
<td>3.91</td>
<td>10.0</td>
<td></td>
</tr>
<tr>
<td>0.05</td>
<td>0.004</td>
<td>0.40</td>
<td>15.6</td>
<td>40.0</td>
<td></td>
</tr>
<tr>
<td>0.01</td>
<td>0.10</td>
<td>10.0</td>
<td>391</td>
<td>1,000.0</td>
<td></td>
</tr>
</tbody>
</table>

\(^30\) For the exercises reported in these tables, the standard deviation of \( \mu_1 \) is, we think, fairly realistically calibrated.
informed and inflation-averse leaders. The reason is that $k > 0$ would be a feature of central bank preferences that accord with a welfare criterion based on household utility when factors such as monopolistic competition or tax distortions, which imply that the flexible-price competitive equilibrium is not socially optimal, are present (see Woodford, 2003). Of course, central banks may regard it as more appropriate to respond to these real distortions using devices other than monetary policy; this, indeed, is the assumption in many recent analyses of optimal monetary policy (including Woodford, 2003, and Aoki, 2001). In any event, knowledge of the relative importance of this bias is relevant for the strategic decisions of central banks.

CONCLUDING REMARKS

We began this analysis by reviewing the distinction between the timeless perspective and discretionary modes of monetary policymaking, the former representing rule-based policy as formalized by Woodford (1999). This distinction becomes important in models with forward-looking expectations, a model feature that was not typically used in the rules-vs.-discretion literature. Typically, there is a second inefficiency from discretionary policymaking, distinct from the more familiar inflationary bias. We calculated the quantitative magnitude of this second inefficiency or loss, using calibrated models of two types prominent in the current literature and a wide range of values representing the relative seriousness of inflation and output-gap variability. The magnitude of the losses is significant, and greater in some (but not all) cases than the inflationary bias arising from a 1 percent excess of the central bank’s output target over the natural-rate value. The losses tend to be somewhat larger in model specifications that imply inflation rate persistence and are often (but not universally) larger with more objective-function weight on output-gap variability.

In addition, we have examined the distinction between instrument rules and targeting rules; our results indicate that targeting-rule outcomes can be closely approximated by instrument rules that respond to any failure of the targeting rule’s optimality condition to hold. Using the instrument rule formulation, we briefly investigated operationality issues, involving the unobservability of current output and perhaps inflation. In addition, we examined a set of cases that assume that the monetary policymaker is using the wrong concept of the natural-rate or potential level of output. In almost all of the various cases examined in the paper, the performance of timeless-perspective policymaking is at least as good as that provided by optimal discretionary behavior. Furthermore, these optimal rules can be well approximated by simple feedback rules based on an interest rate instrument.

REFERENCES


