
MARC P. GIANNONI
MICHAEL WOODFORD

How Forward-Looking is Optimal Monetary Policy?

We calculate optimal monetary policy rules for several variants of a simple optimizing model of the monetary transmission mechanism with sticky prices and/or wages. We show that robustly optimal rules can be represented by interest-rate feedback rules that generalize the celebrated proposal of Taylor (1993). Optimal rules, however, require that the current interest-rate operating target depend positively on the recent past level of the operating target, and its recent rate of increase, in a way that is characteristic of estimated central bank reaction functions, but not of Taylor's proposal. We furthermore find that a robustly optimal policy rule is almost inevitably an *implicit* rule that requires the central bank to use a structural model to project the economy's evolution under the contemplated policy action. However, calibrated examples suggest that optimal rules place less weight on projections of inflation or output many quarters in the future than do rules often discussed in the literature on inflation targeting, or in the current practice of inflation-forecast targeting central banks.

BOTH POSITIVE AND NORMATIVE accounts of monetary policy are often expressed in terms of systematic rules for determining the central bank's operating target for a short-term nominal interest rate in the light of current macroeconomic conditions, especially following the widely discussed proposal of Taylor (1993). Empirically estimated central bank reaction functions are typically similar in form to the Taylor rule, but incorporate additional dynamics. For example, estimated reaction functions, such as those of Judd and Rudebusch (1998), Clarida, Gali, and Gertler (2000), or Nelson (2001), frequently imply that one or more recent past levels of the interest-rate target are important determinants of its current value, along with macroeconomic indicators such as an inflation rate or a measure of

This paper is excerpted from a longer working paper, circulated as "Optimal Interest Rate Rules: II. Applications," NBER working paper no. 9420, December 2002. We thank Tim Fuerst, Jeff Fuhrer, Ed Nelson, Julio Rotemberg, and Lars Svensson for helpful comments, and the National Science Foundation, through a grant to the NBER, for research support.

MARC P. GIANNONI is assistant professor of economics and finance at Columbia University, Graduate School of Business. E-mail: mg2190@columbia.edu MICHAEL WOODFORD is the Harold H. Helm '20 Professor of Economics and Banking, Princeton University. E-mail: woodford@princeton.edu

Journal of Money, Credit, and Banking, Vol. 35, No. 6 (December 2003, Part 2)
Published in 2003 by The Ohio State University Press.

the output gap. This implies that the bank's interest-rate target at any given time depends on past conditions as well as those measured at that time. Furthermore, some estimated reaction functions, such as those of Clarida et al. and Nelson just mentioned, imply that the central bank responds to forecasts of future inflation and/or output rather than to the current values of those variables, as proposed by Taylor. Indeed, the representations of policy incorporated into the econometric models used by some central banks for policy simulations imply that the bank implements a forward-looking Taylor rule, perhaps with a forecast horizon as far as two years in the future.¹ Furthermore, the official explanations that inflation-forecast targeting central banks offer for their decisions, such as in the Inflation Reports of the Bank of England, typically emphasize inflation forecasts with a horizon of that length.²

This paper considers the optimal choice of an interest-rate rule as a basis for the conduct of monetary policy. We shall be particularly concerned with the question of the extent to which it is desirable for the policy rule to be forward looking (as has been emphasized in particular by the literature on inflation targeting), as opposed to seeking to establish a purely contemporaneous relation, as in Taylor's proposal, or even a backward-looking relation of the kind indicated by many econometrically estimated reaction functions.

The question has been extensively discussed in recent years. However, most of the recent literature assumes some low-dimensional parametric family of policy rules, and then optimizes over the coefficients of the rule, using an economic model to compute the equilibrium associated with each possible set of parameters.³ A characteristic weakness of such work,⁴ in our view, is that the conclusions reached about the optimal values of certain parameters are likely to be strongly influenced by the parametric family of rules considered, i.e., by which other kinds of feedback are assumed *not* to be possible.

Hence, we take another approach here, closer to that usually taken in the literature on optimal fiscal policy. We first characterize the optimal state-contingent evolution of the economy, optimizing over the set of stochastic processes consistent with the structural equations that define rational-expectations equilibrium, and then derive a policy rule that can implement the desired equilibrium. In order for the policy rule to "implement" the equilibrium, we not only mean that it must be consistent with it but also that the rule (in conjunction with the structural relations) must determine a *unique* nonexplosive rational-expectations equilibrium. This means that it does not suffice to characterize optimal policy by computing the state-contingent instrument path associated with the optimal equilibrium.⁵ Instead, the optimal policy rule must be described by a relation between endogenous variables that the central bank is committed to bring about, in the spirit of the Taylor rule.

Of course, once we characterize policy by a relation among endogenous variables, there will be many such rules that are each consistent with the desired equilibrium. We therefore obtain a more precise policy recommendation by demanding that our optimal policy rule be *robustly optimal* in the sense discussed in Giannoni and Woodford (2002, section 4): we demand that the rule determine an optimal equilibrium regardless of the assumed statistical properties of the exogenous disturbances.

Such robustness is highly desirable, as it greatly increases the plausibility of a central bank's being willing to commit itself to conduct policy in accordance with the rule. In practice, economies are affected by a large number of different types of disturbances that differ in their expected degree of persistence, the degree to which their effects are expected to be delayed, and so on; and central banks always have a great deal of information about the specific disturbances that have recently hit the economy, even if it would be hard to enumerate all of the possible disturbances in advance. A rule with a separate proviso dealing with each of the types of shocks that could ever occur would not be practical to discuss; but a rule that is optimal only on the assumption that no disturbances will ever occur other than those of a few "typical" types is not one that any central bank is likely to commit itself to follow. A robustly optimal rule eliminates this practical difficulty with the notion of commitment to a policy rule.

The requirement of robust optimality is especially important for clarification of the advantages of forward-looking rules. In the context of a given specification of the statistical properties of the disturbances, and the associated optimal equilibrium, we may find a forward-looking policy rule that is consistent with the equilibrium; but there will necessarily also be a rule that makes no reference to expectations (and that may instead depend on lagged endogenous variables) that is equally consistent with the optimal equilibrium, obtained by replacing the expectation terms in the policy rule by the functions of current and lagged variables that represent rational forecasts in the context of this equilibrium. It is only if we ask whether the same policy continues to be optimal when we vary the statistical properties of the disturbances that we can hope to find an advantage of one representation of the policy rule over the other.

In a previous paper (Giannoni and Woodford 2002), we have expounded a general approach to the design of robustly optimal policy rules, in the context of a fairly general linear-quadratic policy problem. Here we consider the implications of our approach in the context of a particular (admittedly stylized) model of the monetary transmission mechanism, or rather a group of related variant models,⁶ giving particular attention to the degree to which robustly optimal rules are forward- or backward looking.

1. A ROBUSTLY OPTIMAL RULE FOR A SIMPLE FORWARD-LOOKING MODEL

We first illustrate our method of constructing robustly optimal policy rules in the context of a basic optimizing model of the monetary transmission mechanism that has also been used for the analysis of optimal monetary policy in papers such as Woodford (1999), Clarida, Gali, and Gertler (1999), and Giannoni (2001).⁷ The model may be reduced to two structural equations

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

$$\pi_t = \kappa x_t + \beta E_t \pi_{t+1} + u_t \quad (2)$$

for the determination of the inflation rate π_t and the output gap x_t , given the central bank's control of its short-term nominal interest-rate instrument i_t and the evolution of the composite exogenous disturbances r_t^n and u_t . Here the output gap is defined relative to an exogenously varying natural rate of output, chosen to correspond to the gap that belongs among the target variables in the central bank's loss function. The "cost-push shock" u_t then represents exogenous variation in the gap between the flexible price equilibrium level of output and this natural rate, due for example to time-varying distortions that alter the degree of inefficiency of the flexible price equilibrium.⁸ The microfoundations for this model imply that $\sigma, \kappa > 0$, and that $0 < \beta < 1$. The unconditional expectation of the natural rate of interest process is given by $E(r^n) = \bar{r} \equiv -\log \beta > 0$, while the cost-push disturbance is normalized to have an unconditional expectation $E(u) = 0$. Otherwise, our theoretical assumptions place no *a priori* restrictions upon the statistical properties of the disturbance processes, and we shall be interested in policy rules that are optimal in the case of a general specification of the additive disturbance processes of the form discussed in Giannoni and Woodford (2002, section 4).

The assumed objective of monetary policy is to minimize the expected value of a loss criterion of the form

$$W = E_0 \left\{ \sum_{t=0}^{\infty} \beta^t L_t \right\}, \quad (3)$$

where the discount factor β is the same as in Constraint (2), and the loss each period is given by

$$L_t = \pi_t^2 + \lambda_x (x_t - x^*)^2 + \lambda_i (i_t - i^*)^2, \quad (4)$$

for certain optimal levels $x^*, i^* \geq 0$ of the output gap and the nominal interest rate, and certain weights $\lambda_x, \lambda_i > 0$. A welfare-theoretic justification is given for this form of loss function in Woodford (2003, chapter 6), where the parameters are related to those of the structural model.⁹ However, our conclusions below are presented in terms of the parameters of the Loss function (4), and are applicable in the case of any loss function of this general form, whether the weights and target values are the ones that can be justified on welfare-theoretic grounds or not. In the numerical results presented below, the model parameters are calibrated as in Woodford (1999). (For convenience, the parameters are reported in Table 1 below.)

As in Woodford (1999), the state-contingent plan that minimizes the Objective (3)–(4) subject to Constraints (1) and (2) satisfies the first-order conditions

$$\pi_t - \beta^{-1} \sigma \Xi_{1t-1} + \Xi_{2t} - \Xi_{2t-1} = 0, \quad (5)$$

$$\lambda_x (x_t - x^*) + \Xi_{1t} - \beta^{-1} \Xi_{1t-1} - \kappa \Xi_{2t} = 0, \quad (6)$$

TABLE 1
 CALIBRATED PARAMETER VALUES FOR THE BASIC NEO-WICKSELLIAN MODEL

<i>Structural Parameters</i>	
β	0.99
σ^{-1}	0.16
κ	0.024
<i>Shock Processes</i>	
$\rho(\rho^*), \rho(u)$	0.35
<i>Loss Function</i>	
λ_x	0.048
λ_i	0.236

$$\lambda_i(i_t - i^*) + \sigma \Xi_{1t} = 0, \tag{7}$$

for each date $t \geq 0$,¹⁰ together with the initial conditions

$$\Xi_{1,-1} = \Xi_{2,-1} = 0. \tag{8}$$

(Here Ξ_{1t} and Ξ_{2t} are the Lagrange multipliers associated with Constraint (1) and (2), respectively.) In the case that a bounded optimal plan exists, it can be described by equations for $\pi_t, x_t, i_t, \Xi_{1t}$, and Ξ_{2t} as linear functions of $\Xi_{1,t-1}$ and $\Xi_{2,t-1}$ together with the current and expected future values of the exogenous disturbances; these linear equations with constant coefficients apply in all periods $t \geq 0$, starting from the initial Condition (8).

It follows from these first-order conditions that in the case of an optimal commitment that has been in force since at least period $t - 2$, it is possible to infer the values of $\Xi_{1,t-1}$ and $\Xi_{2,t-1}$ from the values that have been observed for x_{t-1}, i_{t-1} , and i_{t-2} . Specifically, one can infer the value of $\Xi_{1,t-1}$ from the value of i_{t-1} using Equation (7), and similarly, the value of $\Xi_{1,t-2}$ from the value of i_{t-2} . Then substituting these values into Equation (6) for period $t - 1$, one can also infer the value of $\Xi_{2,t-1}$ from the value of x_{t-1} . One can, of course, similarly solve for the period t Lagrange multipliers as functions of x_t, i_t , and i_{t-1} . Using these expressions to substitute out the Lagrange multipliers in Equation (5), one obtains a linear relation among the endogenous variables $\pi_t, x_t, x_{t-1}, i_t, i_{t-1}$, and i_{t-2} that must hold in any period $t \geq 2$. Thus, this provides a candidate policy rule that is consistent with the optimal state-contingent plan.

Because the relation in question involves a nonzero coefficient on i_t , it can be expressed as an implicit instrument rule of the form

$$i_t = (1 - \rho_1)i_t^* + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_\pi \pi_t + \phi_x \Delta x_t / 4, \tag{9}$$

where

$$\rho_1 = 1 + \frac{\kappa \sigma}{\beta} > 1, \rho_2 = \beta^{-1} > 1, \tag{10}$$

$$\phi_\pi = \frac{\kappa \sigma}{\lambda_i} > 0, \phi_x = \frac{4 \sigma \lambda_x}{\lambda_i} > 0. \tag{11}$$

We can furthermore show (see Appendix for proof) that commitment to this rule implies a determinate equilibrium.

PROPOSITION 1: *Suppose that a bounded optimal state-contingent plan exists. Then in the case of any parameter values σ , κ , λ_x , $\lambda_i > 0$ and $0 < \beta < 1$, a commitment to the rule described by Rule (9) and Equations (10) and (11) implies a determinate rational-expectations equilibrium.*

The equilibrium determined by commitment to this rule from date $t = 0$ onward corresponds to the unique bounded solution to Equations (5)–(7) when the initial Condition (8) is replaced by the values of $\Xi_{1,-1}$ and $\Xi_{2,-1}$ that would be inferred from the historical values of x_{-1} , i_{-1} , and i_{-2} under the reasoning described above.

It follows that the equilibrium determined by commitment to the time-invariant instrument Rule (9) involves the same responses to random shocks in periods $t \geq 0$ as under the optimal commitment. This is thus an example of an instrument rule that is *optimal from a timeless perspective*, in the sense defined in Giannoni and Woodford (2002, section 3). Note that we could instead implement precisely the optimal once-and-for-all commitment from date $t = 0$ onward (the bounded solution to Equations (5)–(7) with initial Condition (8)) by committing to Rule (9) in all periods $t \geq 2$, but to a modified version of the rule in periods $t = 0$ and 1. But this would be a non-time-invariant rule (policy would depend upon the date relative to the date at which the commitment had been made), and the preferability of this alternative equilibrium, from the standpoint of expected welfare looking forward from date $t = 0$, would result from the alternative policy's optimal exploitation of prior expectations that are already given in that period. Choice of a rule that is optimal from a timeless perspective requires us to instead commit to set the interest rate according to the time-invariant Rule (9) in all periods.

The Rule (9) has the additional advantage of being robustly optimal, in the sense defined in Giannoni and Woodford (2002, section 4). We note that our derivation of the optimal rule has required no hypotheses about the nature of the disturbance processes $\{r_t^n, u_t\}$, except that they are exogenously given and that they are bounded. In fact, the rule is optimal regardless of their nature; commitment to this rule implies the optimal impulse responses displayed in Woodford (1999) in the case of the particular disturbance processes assumed in the numerical illustrations there, but it equally implies optimal responses in the case of any other types of disturbances to the natural rate of interest and/or “cost-push shocks”—disturbances that may be anticipated some quarters in advance, disturbances the effects of which do not die out monotonically with time, and so on.¹¹ Indeed, one may assume that both of the disturbances r_t^n and u_t in Constraints (1) and (2) are composite disturbances of the general form discussed in Giannoni and Woodford (2002, section 4), and Rule (9) remains an optimal rule. This robustness of the rule is a strong advantage from the point of view of its adoption as a practical guide to the conduct of monetary policy.

The optimal Rule (9) has a number of important similarities to the Taylor rule. Like the Taylor rule, Rule (9) is an example of a direct, implicit instrument rule. The rule is also similar to Taylor's recommendation in that the contemporaneous

effect of an increase in either inflation or the output gap upon the federal funds rate operating target is positive ($\phi_\pi, \phi_x > 0$); and the rule satisfies the “Taylor principle,” given that $\phi_\pi > 0$ and $\rho_1 > 1$.¹² However, this optimal rule involves additional history dependence, owing to the nonzero weights on the lagged funds rate, the lagged rate of increase in the funds rate, and the lagged output gap.¹³ And the optimal degree of history dependence is nontrivial: the optimal values of ρ_1 and ρ_2 are both necessarily greater than one, while the optimal coefficient on x_{t-1} is as large (in absolute value) as the coefficient on x_t . It is particularly worth noting that the optimal rule implies not only intrinsic inertia in the dynamics of the funds rate—a transitory deviation of the inflation rate from its average value increases the funds rate not only in the current quarter, but in subsequent quarters as well—but is actually *superinertial*: the implied dynamics for the funds rate are explosive, if the initial overshooting of the long-run average inflation rate is not offset by a subsequent undershooting (as actually always happens, in equilibrium). In this respect this optimal rule is similar to those found to be optimal in the numerical analysis by Rotemberg and Woodford (1999) of a more complicated empirical version of the model.

In the case of the calibrated parameter values in Table 1, the coefficients of the optimal instrument rule are given by $\rho_1 = 1.15$, $\rho_2 = 1.01$, $\phi_\pi = 0.64$, and $\phi_x = 0.33$. These may be compared with the coefficients of the Fed reaction function of similar form estimated by Judd and Rudebusch (1998) for the Greenspan period: $\rho_1 = 0.73$, $\rho_2 = 0.43$, $\phi_\pi = 0.42$, and $\phi_x = 0.30$, except that in this empirical reaction function ϕ_x represents the reaction to the current quarter’s *level* of the output gap, rather than its first difference.¹⁴ (Interestingly, they find that an equation with feedback from the first difference of the output gap, rather than its level, fits best during an earlier period of Fed policy, under Paul Volcker’s chairmanship.) The signs of the coefficients of the optimal rule agree with those characterizing actual policy; in particular, the estimated reaction function includes substantial positive coefficients ρ_1 and ρ_2 , though these are still not as large as the optimal values. Thus, the way in which actual Fed policy is more complex than adherence to a simple Taylor rule can largely be justified as movement in the direction of optimal policy, according to the simple model of the transmission mechanism assumed here.

We find that in the case of this simple model at least, it is not necessary for the central bank’s operating target for the overnight interest rate to respond to forecasts of the future evolution of inflation or of the output gap in order for policy to be fully optimal—and not just optimal in the case of particular assumed stochastic processes for the disturbances, but robustly optimal.¹⁵ Thus, the mere fact that the central bank may sometimes have information about future disturbances, that are not in any way disturbing demand or supply conditions yet, is not a reason for feedback from current and past values of the target variables to be insufficient as a basis for optimal policy. This does not mean that it may not be desirable for monetary policy to restrain spending and/or price increases even before the anticipated real disturbances actually take effect. But in the context of a forward-looking model of

private sector behavior, a commitment to respond to fluctuations in the target variables only contemporaneously and later does not preclude effective preemptive constraint of that kind. First of all, such a policy may well mean that the central bank *does* adjust its policy instrument immediately in response to the news, insofar as forward-looking private sector behavior may result in an immediate effect of the news upon current inflation and output.¹⁶ And more importantly, in the presence of forward-looking private sector behavior, the central bank mainly affects the economy through changes in expectations about the *future* path of its instrument in any event; a predictable adjustment of interest rates *later*, once the disturbances substantially affect inflation and output, should be just as effective in restraining private sector spending and pricing decisions as a preemptive increase in overnight interest rates immediately.

At the same time, it is important to note that the optimal Rule (9), while not “forecast-based” in the sense in which this term is usually understood, *does* depend upon projections of inflation and output in the same quarter as the one for which the operating target is being set. Thus, the rule is not an explicit instrument rule in the sense of Svensson and Woodford (2004). And this implicit character (a feature that it shares with the Taylor rule) is crucial to the optimality of the rule, at least if we wish to find an optimal rule that is also a direct rule (specifying feedback only from the target variables). Optimal policy must generally involve an immediate adjustment of the short-term nominal interest rate in response to shocks, as shown in Woodford (1999);¹⁷ and so unless the rule is to be specified in terms of the central bank’s response to particular shocks, it will have to specify a contemporaneous response to fluctuations in the target variables, and not simply a lagged response. Thus, implementation of such a rule will involve judgment of some sophistication about current conditions; it cannot be implemented mechanically on the basis of a small number of publicly available statistics.

2. OPTIMAL RULES FOR A MODEL WITH INFLATION INERTIA

The basic model considered above is often criticized as being excessively forward looking, particularly in its neglect of any sources of intrinsic inertia in the dynamics of inflation. It might be suspected that this feature of the model is responsible for our strong conclusion above, according to which a robustly optimal policy rule need involve no dependence upon forecasts of the target variables beyond the current period. In Svensson’s (1997) classic argument for the optimality of inflation-forecast targeting, it is the existence of lags in the effect of monetary policy on inflation that causes the optimal rule to involve a target criterion for a forecast, with the optimal forecast horizon coinciding with the length of the policy-transmission lag. It might reasonably be suspected that forecasts are not necessary in our analysis above because our simple model includes no lags in the effects of policy.

Here we take up this question by extending our analysis to the case of a model that incorporates inflation inertia through a device proposed by Christiano, Eichenbaum, and Evans (2001).¹⁸ In this extension of our basic model, prices are not held

constant between the dates at which they are reoptimized, but instead are automatically adjusted on the basis of the most recent quarter’s increase in the aggregate price index, by a percentage that is a fraction γ of the percentage increase in the index. As shown in Woodford (2003, chapter 3), the aggregate supply Relation (2) then takes the more general form

$$\pi_t - \gamma\pi_{t-1} = \kappa x_t + \beta E_t(\pi_{t+1} - \gamma\pi_t) + u_t, \tag{12}$$

where the coefficient κ and the disturbance u_t are defined as before. For γ substantially greater than zero, this makes past inflation an important determinant of current inflation, along with current and expected future output gaps and cost-push shocks; if γ is close enough to one, even a monetary disturbance that has only a transitory effect on real activity can have a much longer-lasting effect on inflation.

The aggregate-demand side of our model remains as before, and our model can accordingly be summarized by the two structural Equations (1) and (12), together with exogenous stochastic processes for the disturbances $\{r_t^r, u_t\}$. As shown in Woodford (2003, chapter 6), the change in our assumptions about pricing behavior implies a corresponding change in the appropriate welfare-theoretic stabilization objective for monetary policy. This is once again a discounted criterion of the Form (3), but the period loss function becomes

$$L_t = (\pi_t - \gamma\pi_{t-1})^2 + \lambda_x(x_t - x^*)^2 + \lambda_i(i_t - i^*)^2. \tag{13}$$

We wish to consider policies that minimize the criterion defined by Equations (3) and (13), subject to the constraints imposed by the structural Equations (1) and (12), for arbitrary values of the indexation parameter $0 \leq \gamma \leq 1$.¹⁹

In the case of this generalization of our policy problem, the first-order Condition (5) becomes instead

$$\begin{aligned} \pi_t^{\text{qd}} - \beta\gamma E_t \pi_{t+1}^{\text{qd}} - \beta^{-1}\sigma \Xi_{1,t-1} - \beta\gamma E_t \Xi_{2,t+1} \\ + (1 + \beta\gamma)\Xi_{2,t} - \Xi_{2,t-1} = 0, \end{aligned} \tag{14}$$

where

$$\pi_t^{\text{qd}} \equiv \pi_t - \gamma\pi_{t-1} \tag{15}$$

is the quasi-differenced inflation rate that appears in both the aggregate supply Relation (12) and the Loss function (13). Conditions (6) and (7) remain as before, and this system of three equations, together with initial Condition (8) and an initial condition for π_{-1} , continues to define the optimal once-and-for-all commitment to apply from date $t = 0$ onward.

As above, we can use Conditions (6) and (7) to substitute for $\Xi_{1,t}$ and $\Xi_{2,t}$ in Equation (14), obtaining an Euler equation of the form

$$E_t[A(L)(i_{t+1} - i^*)] = -f_t \tag{16}$$

for the optimal evolution of the target variables. Here $A(L)$ is a cubic lag polynomial

$$A(L) \equiv \beta\gamma - (1 + \gamma + \beta\gamma)L + (1 + \gamma + \beta^{-1}(1 + \kappa\sigma))L^2 - \beta^{-1}L^3, \quad (17)$$

while the term f_t is a function of the observed and expected future paths of the target variables, defined by

$$f_t \equiv \tilde{q}_t - \beta\gamma E_t \tilde{q}_{t+1}, \quad (18)$$

$$\tilde{q}_t \equiv \frac{\kappa\sigma}{\lambda_i} \left[\pi_t^{\text{gd}} + \frac{\lambda_x}{\kappa} \Delta x_t \right]. \quad (19)$$

By an argument directly analogous to the proof of Proposition 1, we can show that if a bounded optimal state-contingent plan exists, the system obtained by adjoining Equation (16) to the structural Equations (1) and (12) implies a determinate rational-expectations equilibrium, in which the responses to exogenous disturbances are the same as under the optimal commitment. (The only difference between this equilibrium and the optimal once-and-for-all commitment just defined relates to the initial conditions, as in our earlier discussion, and once again this difference is irrelevant to the design of a policy rule that is optimal from a timeless perspective.) Hence, we could regard Equation (16) as implicitly defining a policy rule, and the rule would once again be robustly optimal. In the limiting case that $\gamma = 0$, Equation (16) ceases to involve any dependence upon $E_t i_{t+1}$, and the proposed rule would coincide with the optimal instrument Rule (9) discussed above.

However, Equation (16) is an even less explicit expression for the central bank's interest-rate policy than the implicit instrument rules considered earlier, for (when $\gamma > 0$) it defines i_t only as a function of $E_t i_{t+1}$. This means that the central bank defines the way in which it is committed to set its instrument only as a function of the way that it expects to act further in the future. This failure to express the rule in "closed form" is especially undesirable from the point of view of our question about the optimal forecast horizon for a monetary policy rule. Expression (16) involves no conditional expectations for variables at dates more than one period in the future. However, this does not really mean that the central bank's forecasts for later dates are irrelevant when setting i_t . For this "rule" directs the bank to set i_t as a function of its forecast of i_{t+1} , and (if the same rule is expected to be used to set i_{t+1}) the bank's forecast at t of i_{t+1} should involve its forecast at t of \tilde{q}_{t+2} . It should also involve its forecast of i_{t+2} , and hence (by similar reasoning) its forecast of \tilde{q}_{t+3} , and so on. Hence, it is more revealing to describe the proposed policy rule in a form that eliminates any reference to the future path of interest rates themselves, and instead refers only to the bank's projections of the future paths of inflation and the output gap.²⁰

To obtain an equivalent policy rule of the desired form, we need to partially "solve forward" Equation (16). This requires factorization of the lag polynomial as

$$A(L) \equiv \beta\gamma(1 - \lambda_1 L)(1 - \lambda_2 L)(1 - \lambda_3 L). \quad (20)$$

We note the following properties of the roots of the associated characteristic equation.

PROPOSITION 2: *Suppose that $\sigma, \kappa > 0$, $0 < \beta < 1$, and $0 < \gamma \leq 1$. Then in the Factorization (20) of the polynomial defined in Equation (17), there is necessarily one real root $0 < \lambda_1 < 1$, and two roots outside the unit circle. The latter two roots are either two real roots $\lambda_3 \geq \lambda_2 > 1$, or a complex pair λ_2, λ_3 of roots with real part greater than 1. Three real roots necessarily exist for all small enough $\gamma > 0$, while a complex pair necessarily exists for all γ close enough to 1.*

(See proof in the Appendix.) We use the conventions in the statement of this proposition in referring to the distinct roots in what follows. It is also useful to rewrite Equation (16) as

$$E_t[A(L)\hat{i}_{t+1}] = -\hat{f}_t, \tag{21}$$

where hats denote the deviations of the original variables from the long-run average values implied by the policy rule (Equation 16), or equivalently, by the optimal commitment.

In the case that three real roots exist, the existence of two distinct roots greater than one allows us two distinct ways of “solving forward,” resulting in two alternative relations,

$$(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_t = (\beta\gamma\lambda_3)^{-1}E_t[(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_t], \tag{22}$$

or

$$(1 - \lambda_1 L)(1 - \lambda_3 L)\hat{i}_t = (\beta\gamma\lambda_2)^{-1}E_t[(1 - \lambda_2^{-1}L^{-1})^{-1}\hat{f}_t]. \tag{23}$$

We can also derive other relations of the same form by taking linear combinations of these ones. Of special interest is the relation

$$\begin{aligned} (1 - \lambda_1 L)\left(1 - \frac{\lambda_2 + \lambda_3}{2}L\right)\hat{i}_t &= \frac{1}{2}(\beta\gamma\lambda_3)^{-1}E_t[(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_t] \\ &+ \frac{1}{2}(\beta\gamma\lambda_2)^{-1}E_t[(1 - \lambda_2^{-1}L^{-1})^{-1}\hat{f}_t]. \end{aligned} \tag{24}$$

Here Relations (22) and (23) are defined (with real-valued coefficients) only in the case that three real roots exist, while Relation (24) can also be derived (and has real coefficients on all leads and lags) in the case that λ_2, λ_3 are a complex pair. Because $|\lambda_2|, |\lambda_3| > 1$, the right-hand side of each of these expressions is well defined and describes a bounded stochastic process in the case of any bounded process $\{\hat{f}_t\}$. (In what follows, we shall refer to the three possible expressions for an optimal instrument rule presented in Relations (22)–(24) as Rule I, Rule II, and Rule III, respectively.)

Each of the Relations (22)–(24) can be solved for \hat{i}_t as a function of two of its own lags and expectations at date t regarding current and future values of \hat{f}_t . These can thus be interpreted as implicit instrument rules, each of which now avoids any direct reference to the planned future path of the central bank’s instrument (though assumptions about future monetary policy will be implicit in the inflation and output-gap forecasts). Each of these policy rules is equivalent to Equation (16) and they are accordingly equivalent to one another, in the following sense.

PROPOSITION 3: *Under the assumptions of Proposition 2, and in the case that the Factorization (20) involves three real roots, a pair of bounded processes $\{\hat{i}_t, \hat{f}_t\}$ satisfy any of the Relations (22), (23), or (24) at all dates $t \geq t_0$ if and only they satisfy Equation (21) at all of those same dates. In the case that a complex pair exists, Relation (24) is again equivalent to Equation (21), in the same sense.*

(See proof in Appendix.) Each of the rules thus represents a feasible specification of monetary policy in the case that its coefficients are real valued, and when this is true it implies equilibrium responses to real disturbances that are those associated with an optimal commitment. Accordingly, each represents an optimal policy rule from a timeless perspective. (Note that although the coefficients differ, these are not really different policies. Proposition 3 implies that they involve identical actions, if the bank expects to follow one of them indefinitely, regardless of the model of the economy used to form the conditional forecasts.)

In the case that three real roots exist, we have a choice of representations of optimal policy in terms of an instrument rule, and this time we cannot choose among them on grounds of simplicity. But Rule I seems particularly appealing in this case. This is the rule (among our three possibilities, or any other linear combinations of these) that puts the least weight on forecasts far in the future. It is proper to ask at what rate the weights on forecasts shrink with the forecast horizon, under the assumption that these shrink as fast as possible consistent with robust optimality of the policy rule, if we wish to determine how much forecast dependence is *necessary* for robust optimality.²¹ This choice is also uniquely desirable in the sense that it remains well defined in the limit as γ approaches zero. In this limit, Rule I reduces to

$$(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_t = \hat{f}_t,$$

which is the optimal instrument Rule (9) derived earlier.²² Instead, in the case of any of the other rules, the coefficients on lagged interest rates become unboundedly large as γ approaches zero. Thus Rule I is clearly the preferable specification of policy in the case of small γ . The desire for a rule that varies continuously with γ , so that uncertainty about the precise value of γ will not imply any great uncertainty about how to proceed, then make Rule I an appealing choice over the entire range of γ for which it is defined.

One might think that the same continuity argument could instead be used to argue for the choice of Rule III in all cases, since this is the only one of our optimal instrument rules that continues to be defined for high values of γ . Yet the instruction to follow Rule I if three real roots exist, but Rule III if there is a complex pair, is also a specification that makes all coefficients of the policy rule continuous functions of γ . The reason is that as γ passes through a critical value $\bar{\gamma}$ at which the real roots of the characteristic equation bifurcate, the two larger real roots, λ_2 and λ_3 , come to exactly equal one another. When $\bar{\gamma}$ is approached from the other direction, the imaginary parts of the complex roots λ_2 and λ_3 approach zero; at the bifurcation point their common real value is the repeated real root obtained as the common limit of the two real roots from the other direction. Hence, when $\gamma = \bar{\gamma}$, Rules I, II, and III are all identical. There is thus no ambiguity about whether Rule I or

Rule III should be applied in this case, and no discontinuity in the coefficients of the recommended rule as γ approaches $\bar{\gamma}$ from either direction. At the same time, this proposal results in a rule that remains well defined as γ approaches zero, and for small $\gamma > 0$ results in a rule that is very close to the one previously recommended for an economy with no inflation inertia.

Each of the Rules I, II, and III can be written in the form

$$i_t = (1 - \rho_1)i_t^* + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_\pi F_t(\pi) + \frac{\phi_x}{4} F_t(x) - \theta_\pi \pi_{t-1} - \frac{\theta_x}{4} x_{t-1}, \tag{25}$$

where we have added the constant terms again to indicate the desired level of interest rates (and not just the interest rate relative to its long-run average level), and where

$F_t(z) \equiv \sum_{j=0}^{\infty} \alpha_{z,j} E_t z_{t+j}$ denotes a linear combination of forecasts of the variable z at

various future horizons, with weights $\{\alpha_{z,j}\}$ normalized to sum to one. This form of rule generalizes the Specification (9) that suffices in the case $\gamma = 0$ in two respects: the interest-rate operating target i_t now depends upon lagged inflation in addition to the lagged variables that mattered before, and it now depends upon forecasts of inflation and the output gap in future periods, and not simply upon the projections of those variables for the current period.

Except in these respects, the coefficients are qualitatively similar to those in Rule (9), as indicated by the following proposition.

PROPOSITION 4: *Under the assumptions of Proposition 2, and a loss function with $\lambda_x, \lambda_\pi > 0$, each of Rules I, II, and III has a representation of the Form (25) for all values of γ for which the rule is well defined, and in this representation,*

$$\rho_1 > 1, \rho_2 > 0,$$

$$0 < \theta_\pi \leq \phi_\pi,$$

and

$$0 < \theta_x = \phi_x.$$

Furthermore, for given values of the other parameters, as $\gamma \rightarrow 0$ (for Rule I) the coefficient θ_π approaches zero, though ϕ_π approaches a positive limit; while as $\gamma \rightarrow 1$ (for Rule III) the coefficients θ_π and ϕ_π approach the same positive limit.

(The proof is again in the Appendix.) It is especially noteworthy that once again the optimal instrument rule is superinertial. We also note that once again what should matter is the projected output gap relative to the previous quarter's output gap, rather than the absolute level of the projected gap; and once again interest rates should be increased if the gap is projected to rise. Once again a higher projected inflation rate implies that the interest rate should be increased; but now the degree to which this is true is lower if recent inflation has been high, and in the extreme case $\gamma = 1$, it is only the projected inflation rate relative to the previous quarter's rate that should matter.

The numerical values of these coefficients are plotted, for alternative values of γ ranging between zero and one, in the various panels of Figure 1, where the assumed values for the other parameters are as in Table 1. For all values $\gamma < \bar{\gamma} = 0.35$, there are three real roots, and for each value of γ the three values corresponding to Rules I, II, and III are each plotted; for $\gamma > \bar{\gamma}$, only Rule III is defined. An interesting feature of these plots is that if one considers the coefficients associated with Rule I for $\gamma \leq \bar{\gamma}$ and Rule III for $\gamma \geq \bar{\gamma}$, one observes that the magnitude of each of the coefficients remains roughly the same, regardless of the assumed value of γ . (The exception is θ_π , which approaches zero for small γ , but becomes a substantial positive coefficient for large γ , as indicated by Proposition 4.)

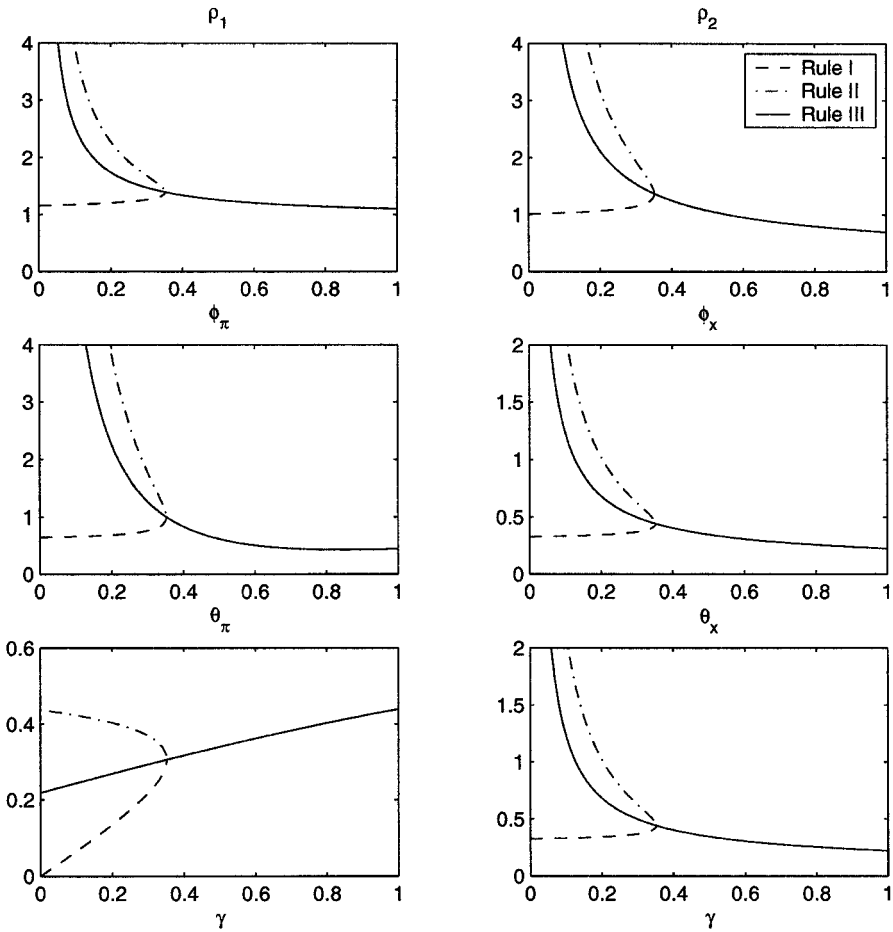


FIG. 1. Coefficients of the optimal instrument Rule (25) as functions of γ

The panels of Figure 2 similarly plot the relative weights $\alpha_{z,j}/\alpha_{z,0}$ for different horizons j of the inflation and output-gap forecasts to which the optimal instrument rule refers,²³ for each of several different possible values of γ . (The weights associated with Rule I are plotted in the case of values $\gamma < \bar{\gamma}$ and those associated with Rule III in the case of values $\gamma > \bar{\gamma}$.) Here we observe that in this case the forecasts $F_t(z)$ are not actually weighted averages of forecasts at different horizons, because the weights are not all nonnegative. Thus while in the presence of inflation inertia, the optimal instrument rule is to some extent forecast based, the optimal responses to forecasts of future inflation and output gaps are not of the sort generally assumed in forward-looking variants of the Taylor rule. In the case of high γ , a higher forecasted inflation rate (or output gap) in any of the next several quarters implies, for given past and projected current conditions, that a *lower* current interest rate is appropriate. According to the optimal rule, a higher current inflation rate should be tolerated in the case that high inflation is forecast for the next several quarters. This is because (in an economy with γ near one) it is sudden *changes* in the inflation rate that creates the greatest distortions in the economy, by making automatic adjustment of prices in response to lagged inflation a poor rule of thumb.

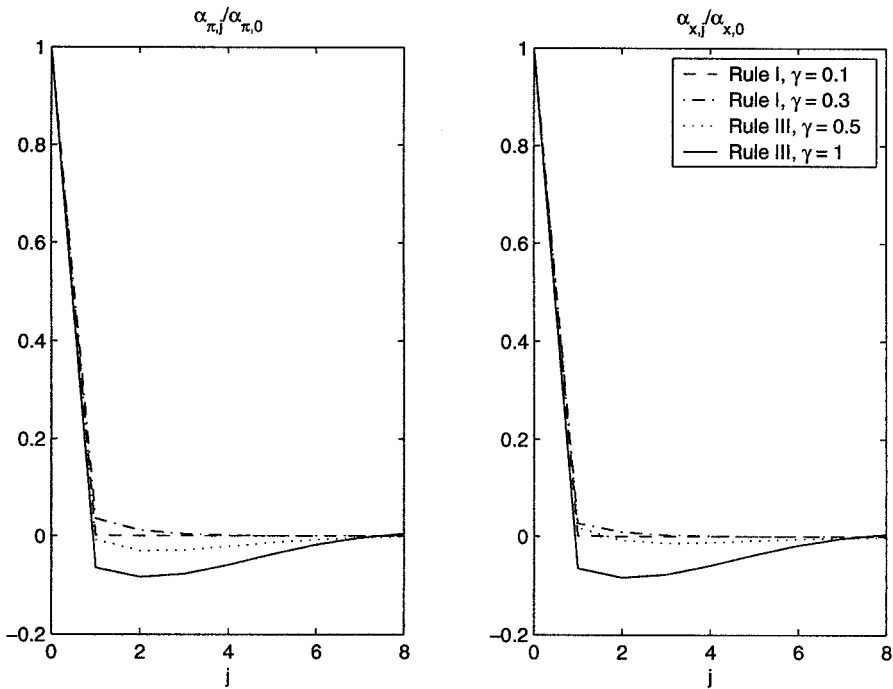


FIG. 2. Relative weights on forecasts at different horizons in the optimal Rule (25)

In addition to this difference from the conventional wisdom with respect to the sign with which forecasts should affect policy, one notes that under the optimal rule it is only forecasts regarding the near future that matter much at all. Even if we consider only the weights put on forecasts for $j \geq 1$ quarters in the future, the mean future horizon of these forecasts, defined by

$$\sum_{j \geq 1} \alpha_{z,j} j / \sum_{j \geq 1} \alpha_{z,j}$$

is equal to only 2.2 quarters in the case of our calibrated example with $\gamma = 1$. Thus, forecasts other than for the first year following the current quarter matter little under the optimal policy. Even more notably, none of the projections beyond the current quarter should receive too great a weight; in our example, the sum of the relative weights on *all* future quarters,

$$\sum_{j > 0} |\alpha_{z,j}| / \alpha_{z,0},$$

is equal to only 0.39 even in the extreme case $\gamma = 1$, while this fraction falls to zero for small γ . Thus, while a robustly optimal direct instrument rule does have to be forecast based in the presence of inflation inertia, the degree to which forecasts matter under the optimal policy rule is still relatively small. Instead, a strong response to projections of inflation and the output gap for the current period, as called for by the Taylor rule, continues to be the crucial element of optimal policy.

3. OPTIMAL POLICY WHEN WAGES ARE ALSO STICKY

As discussed in Woodford (2003, chapter 3), a more realistic model will allow for sticky wages as well as prices. It is argued in this that estimated impulse responses to an identified monetary policy shock are best fit by a model under which wages and prices are roughly equally sticky, so that wages and prices respond to a similar extent (and with similar speed) to a monetary disturbance. Allowing for wages as well as prices to be sticky also creates a further reason for a lagged endogenous variable to matter for inflation determination. For the level of the real wage will be sticky in such a model, and this gradually changing variable is an important determinant of inflation, as a result of its consequences for the real marginal cost of production. The stickiness of wages as well as prices is thus another source of inertia in inflation determination and as such might be expected to justify a more forward-looking policy rule.

Here we consider how the form of an optimal policy rule changes in the case that wages and prices are both sticky to a similar extent. We assume a structural model with monopolistic competition among the suppliers of differentiated types of labor and Calvo-style staggering of wage adjustment, as in Erceg, Henderson, and Levin (2000), and utility-based stabilization objectives similar to the ones derived by these authors.²⁴ For the sake of brevity, we proceed directly to the case of a model that generalizes theirs, discussed in Woodford (2003, chapter 3), in

which both wages and prices are partially indexed to lagged inflation in the way that prices are indexed in the previous section.

The structural equations of our model are²⁵

$$\pi_t - \gamma_p \pi_{t-1} = \kappa_p(x_t + u_t) + \xi_p(\hat{w}_t - \hat{w}_t^j) + \beta E_t[\pi_{t+1} - \gamma_p \pi_t], \quad (26)$$

$$\pi_{w,t} - \gamma_w \pi_{t-1} = \kappa_w(x_t + u_t) + \xi_w(\hat{w}_t^j - \hat{w}_t) + \beta E_t[\pi_{w,t+1} - \gamma_w \pi_t], \quad (27)$$

together with the intertemporal IS relation Equation (1). Here $\pi_{w,t}$ represents nominal wage inflation, \hat{w}_t is the deviation of the log real wage from its steady-state level, \hat{w}_t^j represents the log deviation of the “natural real wage”—i.e., the equilibrium real wage in the case of complete wage and price flexibility—from its steady-state level, and the coefficients $\xi_p, \xi_w, \kappa_p, \kappa_w$ are all positive. The coefficients $0 \leq \gamma_p \leq 1$ and $0 \leq \gamma_w \leq 1$ indicate the degree of indexation of prices and wages, respectively, to the lagged price index, analogous to the indexation of prices in the model of Section 2. (The model of Erceg, Henderson, and Levin (2000) corresponds to the special case in which $\gamma_w = \gamma_p = 0$.)

Under the microeconomic foundations for these relations discussed in Woodford (2003, chapter 3), the appropriate welfare-theoretic stabilization objective is a discounted criterion of Form (3), with a period loss function of the form²⁶

$$L_t = \lambda_p(\pi_t - \gamma_p \pi_{t-1})^2 + \lambda_w(\pi_{w,t} - \gamma_w \pi_{t-1})^2 + \lambda_x(x_t - x^*)^2 + \lambda_i(i_t - i^*)^2. \quad (28)$$

We wish to consider policies that minimize the criterion defined by Equations (3) and (28), subject to the constraints imposed by the structural Equations (1), (26), and (27).

Using the same Lagrangian method as before to characterize optimal policy, we obtain a set of first-order conditions

$$\begin{aligned} & \lambda_p[(\pi_t - \gamma_p \pi_{t-1}) - \beta \gamma_p E_t(\pi_{t+1} - \gamma_p \pi_t)] \\ & - \lambda_w \beta \gamma_w E_t(\pi_{w,t+1} - \gamma_w \pi_t) + \beta^{-1} \sigma \Xi_{1,t-1} \\ & + \beta \gamma_p E_t(\Xi_{2,t+1} - \Xi_{2t}) - (\Xi_{2t} - \Xi_{2,t-1}) \\ & + \beta \gamma_w E_t(\Xi_{3,t+1} - \Xi_{3t}) - \Xi_{4t} = 0, \end{aligned} \quad (29)$$

$$\lambda_w(\pi_{w,t} - \gamma_w \pi_{t-1}) - (\Xi_{3t} - \Xi_{3,t-1}) + \Xi_{4t} = 0, \quad (30)$$

$$\lambda_x(x_t - x^*) - \Xi_{1t} + \beta^{-1} \Xi_{1,t-1} + \kappa_p \Xi_{2t} + \kappa_w \Xi_{3t} = 0, \quad (31)$$

$$\lambda_i(i_t - i^*) - \sigma \Xi_{1t} = 0, \quad (32)$$

$$\xi_p \Xi_{2t} - \xi_w \Xi_{3t} - \Xi_{4t} + \beta E_t \Xi_{4,t+1} = 0, \quad (33)$$

where $\Xi_{1t}, \Xi_{2t}, \Xi_{3t}$ are the Lagrange multipliers associated with Constraints (1), (26), and (27) respectively, and Ξ_{4t} is the multiplier associated with the constraint

$$\hat{w}_t = \hat{w}_{t-1} + \pi_{wt} - \pi_t .$$

A case in which these equations are especially easy to interpret is the special case mentioned earlier, in which $\kappa_w = \kappa_p = \kappa$ (so that wages and prices are sticky to a similar degree) and $\gamma_w = \gamma_p = \gamma$ (so that wages and prices are indexed to lagged inflation to the same degree). In this case, we can add Equation (29) to Equation (30), use Equation (31) to substitute for $\Xi_{2t} + \Xi_{3t}$ and Equation (32) to substitute for Ξ_{1t} , and finally obtain

$$\begin{aligned} & (\lambda_p \pi_t + \lambda_w \pi_{wt} - \gamma \pi_{t-1}) - \beta \gamma E_t (\lambda_p \pi_{t+1} + \lambda_w \pi_{w,t+1} - \gamma \pi_t) \\ &= \frac{\lambda_i}{\kappa \sigma} (\Delta i_t - \beta^{-1} \Delta i_{t-1}) - \beta \gamma E_t \left[\frac{\lambda_i}{\kappa \sigma} (\Delta i_{t+1} - \beta^{-1} \Delta i_t) \right] \\ & \quad - \frac{\lambda_x}{\beta} (i_{t-1} - i^*) - \frac{\lambda_x}{\kappa} (\Delta x_t - \beta \gamma E_t \Delta x_{t+1}) . \end{aligned}$$

This is again an Euler equation of the Form (16), where again $A(L)$ is defined by Equation (17) and f_t is defined by Equations (18) and (19); the only difference is that in the last of these equations, π_t^{qd} is now defined as

$$\pi_t^{\text{qd}} \equiv \lambda_p \pi_t + \lambda_w \pi_{wt} - \gamma \pi_{t-1} \tag{34}$$

rather than as in Equation (15). It follows that optimal policy rules are of essentially the same form as for the model with only sticky prices, except that terms that previously involved only price inflation will now involve both wage and price inflation.

In the case that $\gamma = 0$ (the model of Erceg, Henderson, and Levin, 2000), we obtain an especially simple result. The optimal instrument rule is again of the Form (9), except that instead of responding to current and lagged price inflation π_t , the rule prescribes a response (with the same coefficients as before) to a weighted average of wage and price inflation,

$$\bar{\pi}_t \equiv \lambda_p \pi_t + \lambda_w \pi_{wt} .$$

For the calibrated parameter values suggested in Table 2 below, this index involves equal weights on wage and price inflation.

It is worth noting that we obtain different coefficients here for the optimal policy rule than in Section 2 *only* because the welfare-theoretic loss function is different in the case that wages as well as prices are sticky. If instead of Equation (28) we were to assume a loss function of the Form (4) with arbitrary weights—a common assumption in non-welfare-theoretic analyses of monetary policy rules—we would again have obtained precisely the same optimal policy rules as in Section 2. (This can be seen from the fact that Equation (34) reduces to Equation (15) if $\lambda_w = 0$.) Thus, sticky wages need not imply any difference in the nature of the tradeoff between inflation and output-gap stabilization available to the central bank; the main significance of wage stickiness is that it makes wage stabilization an appropriate

TABLE 2
PARAMETER VALUES IN A CALIBRATED MODEL WITH STICKY WAGES AND PRICES

<i>Structural Parameters</i>	
α_w	0.66
α_p	0.62
β	0.99
σ^{-1}	0.16
ω_w	0.28
ω_p	0.43
$\theta_w \phi^{-1}, \theta_p$	7.88
ξ_w, ξ_p	0.055
κ_w, κ_p	0.024
<i>Loss Function</i>	
λ_w, λ_p	0.50
λ_x	0.048
λ_i	0.236

objective for policy, with consequences for the form of inflation index that belongs in an optimal policy rule.

In the case of indexation to lagged inflation, the roots of the lag polynomial $A(L)$ are the same as in the previous section, yielding the same forms as before for alternative optimal policy rules. (The three optimal instrument rules are each defined for the same values of γ as above; the unique optimal targeting rule is again defined for all γ .) Each of the three optimal instrument rules can be written in the form

$$i_t = (1 - \rho_1)i_t^* + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_p F_t(\pi) + \phi_w F_t(\pi_w) + \frac{\phi_x}{4} F_t(x) - \theta_\pi \pi_{t-1} - \frac{\theta_x}{4} x_{t-1}, \tag{35}$$

where the coefficients $\rho_1, \rho_2, \phi_x, \theta_\pi,$ and θ_x and the coefficients $\{\alpha_{\pi,j}\}$ are all the same functions of the model parameters as in Equation (25), for each of the three rules. The coefficients multiplying the price and wage inflation forecasts satisfy

$$\phi_p + \phi_w = \phi_\pi,$$

and

$$\phi_p \alpha_{p,j} + \phi_w \alpha_{w,j} = \phi_\pi \alpha_{\pi,j}$$

for each $j \geq 0$, where ϕ_π and the $\{\alpha_{\pi,j}\}$ are the coefficients multiplying the inflation forecasts in Equation (25). Thus, if wages and prices are forecasted to increase at the same rate, the effect of these inflation forecasts on the desired interest-rate setting is the same as before. However, if the wage and price inflation forecasts differ, optimal policy now depends on the wage inflation forecast as well.

When $\gamma > 0$, the optimal rule no longer involves only projections of a single-weighted average of wage and price inflation; this is because both wages and prices are (by assumption) indexed only to lagged price inflation, and not to lagged wage

inflation. The difference in the optimal responses to wage inflation as opposed to price inflation is illustrated in Figures 3 and 4 for a calibrated example, with parameter values displayed in Table 2. Here we assume that $\xi_w = \xi_p$, $\kappa_w = \kappa_p$, and $\theta_w \phi^{-1} = \theta_p$. The values assumed for β , σ , θ_p , and κ_p are taken from the estimates of the IS equation and price inflation equation by Rotemberg and Woodford (1997);²⁷ the value of ξ_p is instead taken from the inflation equation estimated by Sbordone (2002), which relates inflation to real marginal cost rather than to the output gap.²⁸ The values of α_w , α_p , ω_w , and ω_p implied by these estimates are also shown in the table,²⁹ as are the implied coefficients λ_w , λ_p , and λ_x of the welfare-theoretic loss function. Note that wage inflation and price inflation receive equal weight, and that the relative weight on output-gap stabilization is the same as in Table 1. We also assume the same relative weight λ_i on interest-rate stabilization as in the calibration above of the flexible-wage model.³⁰

Figure 3 shows the value of the coefficients ϕ_p and ϕ_w in the optimal instrument Rule (35), for alternative values of γ ranging between zero and one. (Again the values are shown for each of Rules I, II, and III in the cases where these exist; the minimally inertial rule corresponds to Rule I for low values of γ and Rule III for high values.) Similarly, Figure 4 shows the relative weights on the inflation and output-gap forecasts at different horizons in the optimal rule.

We note that for moderate positive values of γ , it continues to be the case that the coefficients on the price-inflation forecasts are quite similar to those on the wage-inflation forecasts; essentially, the coefficient ϕ_π shown in Figure 1 is split roughly equally between the coefficients ϕ_p and ϕ_w , while the relative weights on forecasts at different horizons remain similar to those shown in Figure 2. Thus, it is not too bad an approximation to optimal policy to choose the same rule as the one described in Section 2, but to respond to forecasts of an index that is an (roughly equally weighted) average of wage and price inflation. For larger values of γ , however, the optimal responses to forecasts of wage and price inflation are substantially different. The optimal value of ϕ_p remains positive, and similar in magnitude to

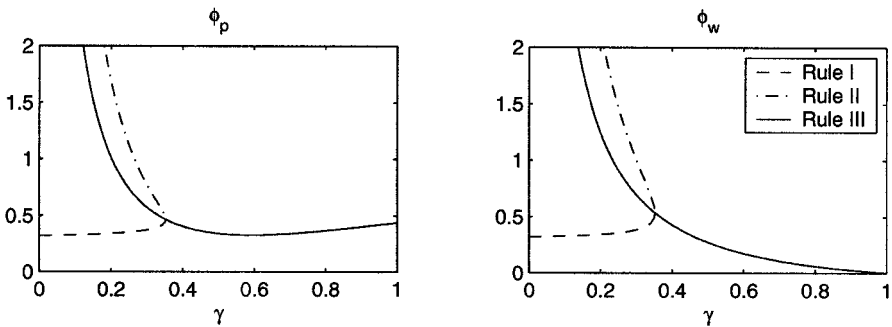


FIG. 3. Coefficients of the optimal instrument Rule (35) as functions of γ . Coefficients not shown are the same as in Figure 1

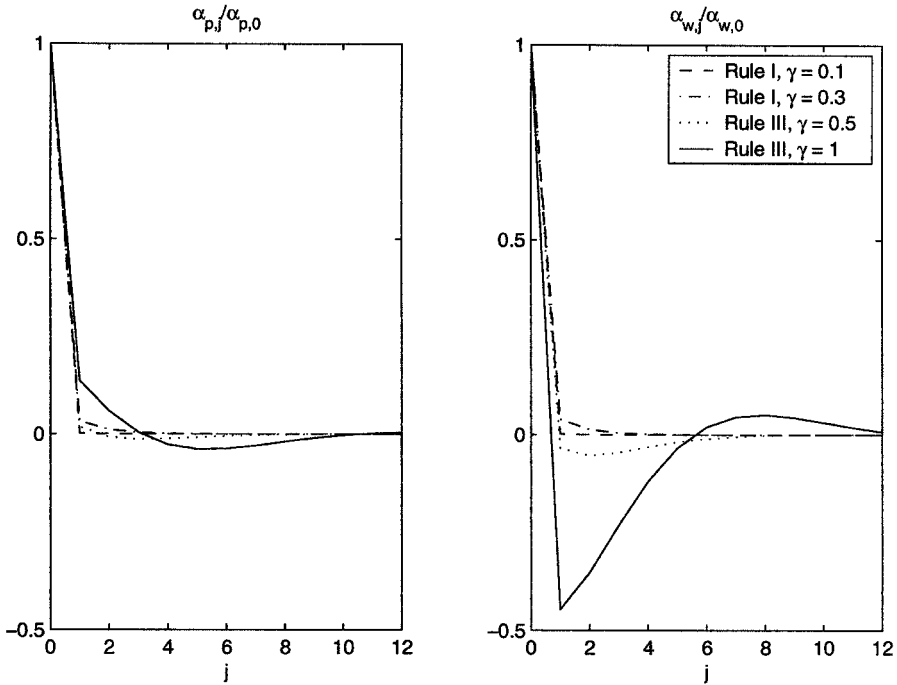


FIG. 4. Relative weights on forecasts at different horizons in the optimal Rule (35)

the previous coefficient ϕ_π , while the optimal value of ϕ_w falls to zero as γ approaches one. This does not mean that it ceases to be optimal to respond to forecasts of wage inflation, only that the sum of the weights at different horizons is zero;³¹ that is, when $\gamma = 1$ the rule prescribes a response only to the forecasted rate of *acceleration* of wage inflation, rather than the rate of wage inflation itself (given the expected rate of price inflation). Specifically, the optimal rule prescribes a negative response to expected deceleration of wage inflation over the next three quarters relative to the current rate of wage inflation; it also prescribes a (weaker) positive response to expected acceleration of wage inflation farther in the future.

Despite these complications, we note that it continues to be the case that optimal policy depends very little on inflation forecasts (either for wages or prices) farther in the future than the coming year, even in the case that there is substantial inflation inertia in both wages and prices. And even with regard to forecasts for the coming year, current interest rates should respond most strongly (and in particular, most *positively*) to projected wage and price inflation in the current quarter, rather than to forecasted inflation later in the year. Thus, there is once again little support for the kinds of forward-looking rules that are sometimes offered as descriptions of the behavior of current inflation-targeting central banks.

4. DELAYS IN THE EFFECTS OF MONETARY POLICY

Empirical models such as those of Rotemberg and Woodford (1997), Amato and Laubach (2003), Christiano, Eichenbaum, and Evans (2001), Altig et al. (2002), or Boivin and Giannoni (2003) differ from the simple models discussed above in that both output and inflation are predetermined, so that neither is immediately affected by an unexpected change in policy. Once again, this is not simply an element the inclusion of which will increase the realism of our model, but one that might be expected to provide stronger justification for forecast-based monetary policy. Here we consider the consequences for optimal policy of allowing for such delays in the effect of policy, modeled in the way described in Woodford (2003, chapter 4, section 4).

Let us consider a model with flexible wages, but sticky prices indexed to lagged inflation, as in Section 2, but now assuming that both price changes and aggregate private demand are predetermined d periods in advance, for some $d \geq 0$. For simplicity, let us suppose that the efficient level of output is also known d periods in advance, so that the output gap is also a predetermined variable.³² In this case, the structural equations of our model are

$$x_t = E_{t-d}x_{t+1} - \sigma E_{t-d}(i_t - \pi_{t+1} - r_t^n), \quad (36)$$

$$\pi_t - \gamma\pi_{t-1} = \kappa E_{t-d}x_t + \beta E_{t-d}(\pi_{t+1} - \gamma\pi_t) + E_{t-d}u_t. \quad (37)$$

The welfare-theoretic loss function continues to be given by Equations (3) and (13). The Lagrangian associated with our policy problem is then of the form

$$\begin{aligned} \mathcal{L} = E_0 \sum_{t=0}^{\infty} \beta^t & \left\{ \frac{1}{2} (\pi_t^{\text{qd}})^2 + \frac{\lambda_x}{2} (x_t - x^*)^2 + \frac{\lambda_i}{2} (i_t - i^*)^2 \right. \\ & + \Xi_{1,t-d} [x_t - E_{t-d}x_{t+1} + \sigma E_{t-d} (i_t - \pi_{t+1} - r_t^n)] \\ & \left. + \Xi_{2,t-d} [\pi_t^{\text{qd}} - \kappa E_{t-d}\pi_{t+1}^{\text{qd}} - E_{t-d}u_t] \right\}, \end{aligned}$$

where π_t^{qd} again denotes the quasi-differenced inflation rate (Equation 15). Here we write $\Xi_{1,t-d}$, $\Xi_{2,t-d}$ for the multipliers associated with Constraints (36) and (37), respectively, to indicate that each multiplier is determined at date $t-d$, given that there is one such constraint for each possible state of the world at date $t-d$.

Using the law of iterated expectations, the Lagrangian can equivalently be written as

$$\begin{aligned} \mathcal{L} = E_0 \sum_{t=d}^{\infty} \beta^t & \left\{ \frac{1}{2} (\pi_t^{\text{qd}})^2 + \frac{\lambda_x}{2} (x_t - x^*)^2 + \frac{\lambda_i}{2} (i_t - i^*)^2 \right. \\ & + \Xi_{1,t-d} [x_t - x_{t+1} + \sigma(i_t - \pi_{t+1})] + \Xi_{2,t-d} [\pi_t^{\text{qd}} - \kappa\pi_{t+1}^{\text{qd}}] \left. \right\} \\ & + \frac{\lambda_i}{2} E_0 \sum_{t=0}^{d-1} \beta^t (i_t - i^*)^2, \quad (38) \end{aligned}$$

dropping terms that are independent of policy. The first-order conditions that characterize an optimal once-and-for-all commitment as of date zero are given by

$$\begin{aligned} \pi_t^{\text{qd}} - \beta\gamma E_{t-d}\pi_{t+1}^{\text{qd}} - \beta^{-1}\sigma\Xi_{1,t-d-1} - \beta\gamma E_{t-d}\Xi_{2,t-d+1} \\ + (1 + \beta\gamma)\Xi_{2,t-d} - \Xi_{2,t-d-1} = 0 \end{aligned} \tag{39}$$

together with Conditions (6) and (7), but with $\Xi_{i,t-d}$ substituted for Ξ_{it} (for $i = 1, 2$) in the latter equations. Each of the first-order conditions just listed holds for each $t \geq d$. In addition, π_{d-1} is given as an initial condition, the initial lagged Lagrange multipliers satisfy Condition (8), and one has additional first-order conditions

$$i_t = i^*$$

for the interest rate in periods $t = 0, \dots, d - 1$. Note, however, that these last conditions, that relate only to the first few periods following the adoption of the optimal commitment, are irrelevant to the characterization of optimal policy from a timeless perspective.

As above, we can use Conditions (6) and (7) to substitute for $\Xi_{1,t-d}$ and $\Xi_{2,t-d}$ in Equation (39), obtaining an Euler equation of the form

$$A_1(L)(i_t - i^*) + \beta\gamma E_{t-d}(i_{t+1} - i^*) = -E_{t-d}f_t, \tag{40}$$

where $A_1(L)$ is the quadratic lag polynomial such that

$$A(L) = \beta\gamma + LA_1(L)$$

is the polynomial defined in Equation (17), and f_t is again defined in Equations (18) and (19). It follows from this that under a policy that is optimal from a timeless perspective, i_t depends solely on public information at date $t - d$. Hence, in the case of structural equations of this kind, there would be no change in the character of optimal policy were one to impose the constraint that the interest-rate operating target must be chosen in advance, as proposed for example by McCallum and Nelson (1999).

Taking the expectation of Equation (40) conditional upon information at date $t - d$, one obtains

$$E_{t-d}[A(L)(i_{t+1} - i^*)] = -E_{t-d}f_t,$$

which is identical to Equation (16) except for the conditioning information set. The same manipulations as before can then be used to derive the same form of representations for optimal policy, with the change in the conditioning information set for expectations. There are once again the three possible forms for an optimal instrument rule discussed in Section 2 above, and each exists for the same values of γ as before. Each of the three rules is of the form

$$\begin{aligned}
i_t = & (1 - \rho_1)i_t^* + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_\pi E_{t-d} F_t(\pi) \\
& + \frac{\phi_x}{4} E_{t-d} F_t(x) - \theta_\pi \pi_{t-1} - \frac{\theta_x}{4} x_{t-1}, \tag{41}
\end{aligned}$$

where the coefficients are exactly the same functions of the model parameters as in Equation (25). Thus, the optimal policy rules are of exactly the same form as before, except that now the period t interest rate should be chosen in period $t - d$, and on the basis of the inflation and output-gap projections that are available at that earlier date. The projections, however, should be for the same time periods as before.

In the case that $\gamma = 0$ (our baseline model, with a standard “New Keynesian Phillips curve,” except for the d - period delays), the projections in Equation (41) again are only for inflation and the output gap in period t . Since both of these variables are known at date $t - d$, according to our model, the optimal instrument rule is once again of the Form (9), with coefficients given in Equations (10) and (11). Thus, in this case there is *no change* at all in the optimal policy rule. It remains true that the delays imply that it is optimal for nominal interest rates to be perfectly forecastable d periods in advance. However, this principle does not imply that a rule that prescribes a response to contemporaneous inflation and output-gap variations, as under the Taylor rule, is therefore suboptimal. For under our assumptions, inflation and the output gap are themselves completely forecastable d periods in advance. This example shows that an optimal policy rule need not be at all forward looking, even in the case that the effects of monetary policy are entirely delayed.

Of course, even if there are no effects of a change in monetary policy until d periods later, it need not follow that inflation and the output gap are completely predetermined. Only the components of these variables that are affected by monetary policy need to be predetermined. In a more complex model, we may assume that the forecastable components of inflation and the output gap, $E_{t-d}\pi_t$ and $E_{t-d}x_t$, satisfy Equations (36) and (37), while the observed variables are equal to these forecastable components plus exogenous disturbance terms.³³ In this case, both Equations (36) and (37) should include additional unforecastable disturbance terms, as in the model discussed in Svensson and Woodford (2004). In this case, the Lagrangian Equation (38) is still correct, up to terms that are independent of policy, and the same first-order conditions continue to apply. The optimal policy rules just derived continue to be correct, except that the central bank should respond only to variation in the *forecastable components* of inflation and the output gap. For example, the optimal instrument rule takes the form

$$\begin{aligned}
i_t = & (1 - \rho_1)i_t^* + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_\pi E_{t-d} F_t(\pi) \\
& + \frac{\phi_x}{4} E_{t-d} F_t(x) - \theta_\pi E_{t-d-1} \pi_{t-1} - \frac{\theta_x}{4} E_{t-d} x_{t-1}
\end{aligned}$$

instead of Equation (41), and even when $\gamma = 0$ the central bank should respond to forecasts of inflation and output d periods in advance, rather than to current inflation and output.

We find that our previous conclusions about the character of optimal policy remain largely intact, even when we allow for delays in the effect of monetary policy. An optimal instrument rule still involves interest-rate inertia to exactly the same degree as was determined earlier; in particular, the optimal rule is superinertial for all possible values of the parameters. We also find once again that optimal policy is only modestly forward looking. If the components of inflation and the output gap that are affected by monetary policy are determined d periods in advance, it follows that policy should respond only to forecasts of inflation and the output gap d or more periods in the future. However, the interest rate in any given period should be set (d periods earlier) on the basis of the projected inflation rate and output gap for the period in which the interest rate applies and periods immediately thereafter; and even when the degree of inflation inertia is substantial, interest rates should be based mainly on projections for that period and a few months farther in the future. There continues to be little support for the idea that primary emphasis should be placed on inflation forecasts for a period one to two years later than the period for which the interest rate is set.

5. CONCLUSIONS

We have shown that robustly optimal policy rules can be constructed for each of a variety of simple forward-looking models of the monetary transmission mechanism. Our results allow us some tentative conclusions about the desirability the kind of delayed response of the level of nominal interest rates to changes in inflation and output that is implied by many estimated central bank reaction functions. Even in our baseline model, which posits an extremely simple dynamic structure, our optimal policy rules involve substantial history dependence of a kind not present in simple proposals such as the Taylor rule. In addition to the fact that policy should respond to the projected *change* in the output gap rather than its level, which makes the recent *past* level of the output gap relevant for current policy, we find that past nominal interest rates should affect the current policy setting. Specifically, our optimal instrument Rule (9) implies that, for any given inflation and output-gap projections, interest rates should be higher than they otherwise would be if (1) interest rates have recently been higher than average or (2) interest rates have recently been rising. Thus, the optimal rules incorporate both the interest-rate *persistence* (a positive effect of i_{t-1} on the choice of i_t) and interest-rate *momentum* (a positive effect of Δi_{t-1} on the choice of Δi_t) that characterize the actual Fed reaction functions estimated by Judd and Rudebusch (1998).

We have also explored the degree to which optimal rules should make policy a function of projections of inflation and/or output many quarters in the future. In our

baseline model, it is possible to formulate a robustly optimal policy rule (the implicit instrument Rule (9) that involves no projections farther in the future than the period for which the nominal interest-rate operating target is being set. Perhaps surprisingly, this rule is optimal regardless of what we may assume about the availability of advance information about future disturbances. Of course, this strong result depends on the purely forward-looking character of that simple model of inflation and output determination. But even when we allow for a high degree of inflation inertia, in Section 2, we find that an optimal policy rule depends much more on the projected inflation rate and output gap in the quarter for which policy is being set than on the projections for any later horizons. And while projections for later quarters do matter to some extent if the degree of inflation inertia is sufficiently great, projections farther than a year in the future matter little even in this case. Thus, we find little justification for a policy that gives primary attention to the inflation forecast at a horizon two years in the future, as is true of the inflation-forecast targeting currently practiced by central banks such as the Bank of England.

It is important nonetheless to stress that our results do not justify a purely backward-looking approach to the conduct of policy. In all of the cases considered, our optimal rules are *implicit* rules, which is to say that they specify a criterion that must be satisfied by the central bank's projections of inflation and output given its policy. The criterion in question involves variables, the values of which depend on the current policy action that is chosen; hence, they must be *projected* using a model of the monetary transmission mechanism, rather than simply being *measured*. It is true that optimal policy could also be described by an explicit (purely backward-looking) instrument rule, specifying the instrument setting as a function of current exogenous disturbances and past (or at any rate predetermined) state variables, that need to be simply measured. But such a representation of optimal policy would not be robust to changes in the assumed character of the disturbance processes, unlike the implicit rules derived here. Hence, we would argue that the use of a quantitative model, that can be used to project the effects of prospective policy settings, is essential to the optimal conduct of monetary policy. And in a model that takes account of forward-looking private sector behavior, projections for the current quarter cannot generally be made without forecasting the economy's subsequent evolution as well.

Furthermore, in the case that spending and pricing decisions are predetermined, as assumed in empirical models such as those of Rotemberg and Woodford (1997), Christiano, Eichenbaum, and Evans (2001), Altig et al. (2002), or Boivin and Giannoni (2003), the optimal policy is one under which the interest-operating target is chosen d periods in advance, on the basis of projections of inflation and output for the period for which the interest rate is being chosen (if not projections farther in the future as well). In this case, policy decisions necessarily will depend crucially on projections of conditions at least d periods in the future. However, the lag d by which spending or pricing decisions are predetermined is not plausibly longer than one or two quarters. And even in this case, no justification is provided for basing the interest-rate operating target for a given period on forecasts regarding points

in time that are much more distant than the period for which the interest-rate decision is being made. Hence, our results provide little support for the desirability of basing interest-rate decisions primarily on forecasts of conditions as long as two years in the future.

APPENDIX A: PROOFS OF PROPOSITIONS

We begin by establishing a technical lemma required in the proof of Proposition 1.

LEMMA 1: *For any real coefficients A_0, A_1, A_2, A_3 , we have*

$$A_3z^3 + A_2z^2 + A_1z + A_0 = A_3(B_0 + B_1\zeta + \zeta^3),$$

where

$$\zeta = z + \frac{A_2}{3A_3},$$

$$B_0 = \frac{27A_0A_3^2 + 2A_2^3 - 9A_3A_2A_1}{27A_3^3},$$

and

$$B_1 = \frac{3A_3A_1 - A_2^2}{3A_3^2}.$$

PROOF: $A_3(B_0 + B_1\zeta + \zeta^3)$

$$= A_3\left(\zeta^3 + \frac{1}{3} \frac{3A_3A_1 - A_2^2}{A_3^2} \zeta + \frac{1}{27} \frac{27A_0A_3^2 + 2A_2^3 - 9A_3A_2A_1}{A_3^3}\right)$$

$$= A_3\left(\left(z + \frac{1}{3} \frac{A_2}{A_3}\right)^3 + \frac{1}{3} \frac{3A_3A_1 - A_2^2}{A_3^2} \left(z + \frac{1}{3} \frac{A_2}{A_3}\right)\right)$$

$$+ \frac{1}{27} \frac{27A_0A_3^2 + 2A_2^3 - 9A_3A_2A_1}{A_3^3}$$

$$= A_3z^3 + A_2z^2 + \left(\frac{1}{3} \frac{A_2^2}{A_3} + \frac{1}{3} \frac{3A_3A_1 - A_2^2}{A_3}\right)z + \frac{1}{27} \frac{A_2^3}{A_3^3}$$

$$+ \frac{1}{27} \frac{27A_0A_3^2 + 2A_2^3 - 9A_3A_2A_1}{A_3^3} + \frac{1}{9} \frac{3A_3A_1 - A_2^2}{A_3^2} A_2$$

$$= A_3z^3 + A_2z^2 + A_1z + A_0. \quad \blacksquare$$

A1. Proposition 1

PROPOSITION 1: *Suppose that a bounded optimal state-contingent plan exists. Then in the case of any parameter values $\sigma, \kappa, \lambda_x, \lambda_i > 0$ and $0 < \beta < 1$, a commitment to the rule described by Rule (9) and Equations (10) and (11) implies a determinate rational-expectations equilibrium.*

PROOF: The system of equations given by the structural Equations (1), (2), and the policy Rule (9) and Equations (10) and (11) can be written in matrix form as

$$\bar{I} \begin{bmatrix} Z_{t+1} \\ E_t z_{t+1} \\ E_t i_{t+1} \end{bmatrix} = \begin{bmatrix} 0 \\ -\bar{\Phi} \end{bmatrix} + \bar{A} \begin{bmatrix} Z_t \\ z_t \\ i_t \end{bmatrix} + \bar{C} s_t, \tag{A1}$$

where $z_t \equiv [\pi_t, x_t]'$, $Z_t \equiv [x_{t-1}, i_{t-1}, i_{t-2}]'$, $s_t \equiv [r_t^n, u_t]'$, and

$$\bar{I} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma & 1 & 0 \\ 0 & 0 & 0 & \beta & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\bar{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & \sigma \\ 0 & 0 & 0 & 0 & 1 & -\kappa \\ -\frac{\sigma\lambda_x}{\lambda_i} & 1 + \frac{\kappa\sigma}{\beta} + \beta^{-1} & -\beta^{-1} & \frac{\kappa\sigma}{\lambda_i} & \frac{\sigma\lambda_x}{\lambda_i} & -1 \end{bmatrix}.$$

Equilibrium is determinate if the characteristic polynomial $\det[\bar{A} - \mu\bar{I}]$ has exactly $n_Z = 3$ roots such that $|\mu| < 1$. Recall that if there are fewer such roots, there is no bounded solution at all. Since Rule (9) and Equations (10) and (11) are derived from the first-order Conditions (5)–(7), it must be consistent with the optimal state-contingent plan. Because we assume that a bounded optimal state-contingent plan exists, it must be the case that $\det[\bar{A} - \mu\bar{I}]$ admits at least three roots inside the unit circle.

Note that we can rewrite the characteristic polynomial as

$$\det[\bar{A} - \mu\bar{I}] = -p(\mu)\beta\mu, \tag{A2}$$

where

$$p(\mu) \equiv \mu^4 - a\mu^3 + b\mu^2 - a\beta^{-1}\mu + \beta^{-2} \tag{A3}$$

and

$$a = 2 \frac{1 + \beta + \sigma\kappa}{\beta} + \frac{\sigma^2 \lambda_x}{\lambda_i},$$

$$b = \frac{1 + 2\kappa\beta\sigma + 2\sigma\kappa + \sigma^2 \kappa^2 + 4\beta + \beta^2}{\beta^2} + \sigma^2 \frac{(1 + \beta)\lambda_x + \kappa^2}{\beta\lambda_i}.$$

We can furthermore express $p(\mu)$ as

$$p(\mu) = (\mu - \mu_1)(\mu - \mu_2)(\mu - \mu_3)(\mu - \mu_4),$$

where, because of the symmetry in Equation (A3), the four roots μ_i satisfy

$$\mu_1 = (\beta\mu_2)^{-1} \text{ and } \mu_3 = (\beta\mu_4)^{-1}. \tag{A4}$$

Because $\det[\bar{A} - \mu\bar{I}]$ admits at least three roots inside the unit circle, Equations (A2) and (A3) imply that $p(\mu)$ admits either two, three, or four roots inside the unit circle. Let us consider each case in turn.

1. Let us suppose first, as a way of contradiction, that all four roots of $p(\mu)$ are inside the unit circle. Then $|\mu_1| < 1$ by assumption. However, Equation (A4) implies $|\beta\mu_2| > 1$, and thus $|\mu_2| > 1$, which contradicts the assumption that all four roots are inside the unit circle.
2. Let us suppose next that $p(\mu)$ has three roots inside the unit circle. If $|\mu_1| < 1$, then Equation (A4) implies again $|\mu_2| > 1$. It follows that the remaining two roots μ_3 and μ_4 must be inside the unit circle. But this is impossible, as $|\mu_3| < 1$ implies $|\mu_4| > 1$. Inversely, if $|\mu_1| > 1$, then the three remaining roots must be inside the unit circle. Again, this is impossible as $|\mu_3| < 1$ implies $|\mu_4| > 1$.

It follows that $p(\mu)$ must have exactly two roots inside the unit circle, and thus that the equilibrium is determinate.

A2. Proposition 2

PROPOSITION 2: *Suppose that $\sigma, \kappa > 0, 0 < \beta < 1$, and $0 < \gamma \leq 1$. Then in the factorization*

$$A(L) = \beta\gamma(1 - \lambda_1 L)(1 - \lambda_2 L)(1 - \lambda_3 L) \tag{A5}$$

of the polynomial

$$A(L) \equiv \beta\gamma - (1 + \gamma + \beta\gamma)L + (1 + \gamma + \beta^{-1}(1 + \kappa\sigma))L^2 - \beta^{-1}L^3, \tag{A6}$$

there is necessarily one real root $0 < \lambda_1 < 1$, and two roots outside the unit circle. The latter two roots are either two real roots $\lambda_3 \geq \lambda_2 > 1$, or a complex pair λ_2, λ_3 with real part greater than 1. Three real roots necessarily exist for all small enough $\gamma > 0$, while a complex pair necessarily exists for all γ close enough to 1.

PROOF: Consider the following properties of the Polynomial (A6):

$$\begin{aligned} A(z) &> 0, \forall z \leq 0 & A'(z) &< 0, \forall z \leq 0 \\ A(\beta) &= \beta\kappa\sigma > 0 & A'(\beta) &= (1 - \beta)(1 - \gamma) + 2\kappa\sigma > 0 \\ A(1) &= \beta^{-1}\kappa\sigma > 0 \\ A(+\infty) &= -\infty \end{aligned}$$

From this, we know that for all $z \leq 0$, $A(z)$ is positive and decreasing. As z is raised from 0 to β , $A(z)$ continues to decrease, reaches a minimum (where $A(z)$ may be positive or negative), and starts increasing as z approaches β . The polynomial $A(z)$ is positive for $z = 1$, but decreases again and tends to be $-\infty$, as z becomes larger and larger. It follows that $A(z)$ admits one real root $z_1 > 1$ and either two real roots $0 < z_3 \leq z_2 < 1$ or a pair of complex roots z_2, z_3 .

Thus, $A(L)$ can be written as

$$\begin{aligned} A(L) &= \beta^{-1}(z_1 - L)(z_2 - L)(z_3 - L) \\ &= \beta^{-1}\lambda_1^{-1}\lambda_2^{-1}\lambda_3^{-1}(1 - \lambda_1 L)(1 - \lambda_2 L)(1 - \lambda_3 L) \\ &= \frac{1}{\beta\lambda_1\lambda_2\lambda_3} - \frac{\lambda_1 + \lambda_2 + \lambda_3}{\beta\lambda_1\lambda_2\lambda_3}L + \frac{\lambda_1\lambda_2 + (\lambda_1 + \lambda_2)\lambda_3}{\beta\lambda_1\lambda_2\lambda_3}L^2 - \beta^{-1}L^3, \end{aligned} \tag{A7}$$

where $\lambda_j \equiv z_j^{-1}$ for $j = 1, 2, 3$. Comparing the first terms of Equations (A7) and (A6), we note that

$$(\beta\lambda_1\lambda_2\lambda_3)^{-1} = \beta\gamma, \tag{A8}$$

so that the polynomial $A(L)$ can be factorized as in Equation (A5), where $0 < \lambda_1 < 1$ and λ_2, λ_3 are either two real roots satisfying $1 < \lambda_2 \leq \lambda_3$, or a pair of complex roots.

We now show that in the case that λ_2, λ_3 form a pair of complex roots, their common real part is greater than 1. Comparing the second term of Equation (A7) with the corresponding term in Equation (A6), and using Equation (A8), we note that

$$\beta\gamma(\lambda_1 + \lambda_2 + \lambda_3) = 1 + \gamma + \beta\gamma. \tag{A9}$$

Furthermore, as $\beta\gamma\lambda_1 < 1$, we have

$$\beta\gamma\lambda_1 = 1 + \gamma - \beta\gamma(\lambda_2 + \lambda_3 - 1) < 1.$$

This implies

$$-\beta\gamma(\lambda_2 + \lambda_3 - 1) < -\gamma,$$

and thus

$$\lambda_2 + \lambda_3 > 1 + \beta^{-1} > 2.$$

Therefore

$$\text{Re}\lambda_2 = \text{Re}\lambda_3 = \frac{\lambda_2 + \lambda_3}{2} > 1.$$

It follows that the moduli $|\lambda_2| = |\lambda_3| > 1$.

We now show that three real roots $\lambda_1, \lambda_2, \lambda_3$ necessarily exist for all small enough $\gamma > 0$, while a complex pair λ_2, λ_3 necessarily exists for all γ close enough to 1. First note that each λ_j , for $j = 1, 2, 3$, is real if and only if the solution $z_j \equiv \lambda_j^{-1}$ of the equation

$$A(z) \equiv A_0 + A_1z + A_2z^2 + A_3z^3 = 0$$

is real, where

$$A_0 = \beta\gamma,$$

$$A_1 = -(1 + \gamma + \beta\gamma),$$

$$A_2 = 1 + \gamma + \beta^{-1}(1 + \kappa\sigma),$$

and

$$A_3 = -\beta^{-1}.$$

Furthermore, since

$$A(z) = A_3(B_0 + B_1\zeta + \zeta^3),$$

where

$$\zeta = z + \frac{A_2}{3A_3}$$

and

$$B_0 = \frac{27A_0A_3^2 + 2A_2^3 - 9A_3A_2A_1}{27A_3^3},$$

$$B_1 = \frac{3A_3A_1 - A_2^2}{3A_3^2}$$

are real coefficients, each λ_j is real if and only if the corresponding solution ζ_j of the equation

$$B_0 + B_1\zeta + \zeta^3 = 0$$

is real (see Lemma 1). From Cardano's formulas for the roots of a cubic equation, we know that this equation admits:

1. three different real roots if $\Delta \equiv 27B_0^2 + 4B_1^3 < 0$,
2. three real roots, at least two of which are equal, if $\Delta = 0$,
3. one real root and two complex roots if $\Delta > 0$.

Expressing Δ as a function of γ , we have:

$$\begin{aligned} \Delta(\gamma) &= 27B_0^2 + 4B_1^3 = A_3^{-4}(27A_0^2A_3^2 + 4A_0A_2^3 - 18A_0A_3A_2A_1 - A_2^2A_1^2 + 4A_3A_1^3) \\ &= -\beta^4(1 - \beta)^2\gamma^4 - 2\beta^3((1 + \beta^2 - 4\beta)\kappa\sigma - (1 + \beta)(1 - \beta)^2)\gamma^3 \\ &\quad - \beta^2(\kappa^2\sigma^2(1 - 10\beta + \beta^2) + 2\kappa\sigma(1 + \beta)(1 + \beta + \beta^2) \\ &\quad + (1 + 4\beta + \beta^2)(1 - \beta)^2)\gamma^2 + \beta^2(4\kappa^3\sigma^3 + 10(1 + \beta)\kappa^2\sigma^2 \\ &\quad + 4(2 + 2\beta^2 - \beta)\kappa\sigma + 2(1 + \beta)(1 - \beta)^2)\gamma \\ &\quad - \beta^2(2\beta\kappa\sigma + \kappa^2\sigma^2 + 2\kappa\sigma + (1 - \beta)^2), \end{aligned}$$

which is a fourth-order polynomial in γ . Note that $\Delta(\gamma)$ is a continuous function of γ that admits at most four real roots and has the following properties:

$$\begin{aligned} \Delta(-\infty) &= -\infty \\ \Delta(0) &= -\beta^2(2\beta\kappa\sigma + \kappa^2\sigma^2 + 2\kappa\sigma + (1 - \beta)^2) < 0 \\ \Delta(1) &= \beta^2\kappa\sigma(4\kappa^2\sigma^2 + (8 + 20\beta - \beta^2)\kappa\sigma + 4(1 - \beta)^3) > 0 \\ \Delta(+\infty) &= -\infty. \end{aligned}$$

It follows that $\Delta(\gamma)$ admits either one or three roots between 0 and 1. Furthermore, $\Delta < 0$ for $\gamma > 0$ small enough, and $\Delta > 0$ for all γ close enough to 1. Thus, three real roots necessarily exist for all small enough $\gamma > 0$, while a complex pair necessarily exists for all γ close enough to 1.

A3. Proposition 3

PROPOSITION 3: *Under the assumptions of Proposition 2, and in the case that the Factorization (20) involves three real roots, a pair of bounded processes $\{\hat{i}_t, \hat{f}_t\}$ satisfies any of the equations*

$$(1 - \lambda_1L)(1 - \lambda_2L)\hat{i}_t = (\beta\gamma\lambda_3)^{-1}E_t[(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_t], \tag{A10}$$

$$(1 - \lambda_1L)(1 - \lambda_3L)\hat{i}_t = (\beta\gamma\lambda_2)^{-1}E_t[(1 - \lambda_2^{-1}L^{-1})^{-1}\hat{f}_t], \tag{A11}$$

or

$$\begin{aligned} (1 - \lambda_1L)(1 - \frac{\lambda_2 + \lambda_3}{2}L)\hat{i}_t &= \frac{1}{2}(\beta\gamma\lambda_3)^{-1}E_t[(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_t] \\ &\quad + \frac{1}{2}(\beta\gamma\lambda_2)^{-1}E_t[(1 - \lambda_2^{-1}L^{-1})^{-1}\hat{f}_t] \end{aligned} \tag{A12}$$

at all dates $t \geq t_0$ if and only they satisfy

$$E_t[A(L)\hat{i}_{t+1}] = -\hat{f}_t \tag{A13}$$

at all of those same dates. In the case that a complex pair exists, Equation (A12) is again equivalent to Equation (A13) in the same sense.

PROOF: Proposition 2 guarantees that the roots $\lambda_1, \lambda_2, \lambda_3$ in Factorization (20) are either real and satisfy $0 < \lambda_1 < 1 < \lambda_2 \leq \lambda_3$ or $0 < \lambda_1 < 1$, and λ_2, λ_3 are complex conjugates that lie outside the unit circle. Consider first the case in which $\lambda_1, \lambda_2, \lambda_3$ are real.

Rule I. We now show that Equation (A13) implies Equation (A10). Using Factorization (20) to substitute for $A(L)$ and expanding the left-hand side of Equation (A13), we obtain

$$\beta\gamma E_t[(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_{t+1} - (1 - \lambda_1 L)(1 - \lambda_2 L)\lambda_3 \hat{f}_t] = -\hat{f}_t$$

or

$$(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_t = (\beta\gamma\lambda_3)^{-1}\hat{f}_t + \lambda_3^{-1}E_t[(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_{t+1}].$$

Substituting recursively for $(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_{t+j}$ on the right-hand side, we obtain

$$\begin{aligned} (1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_t &= (\beta\gamma\lambda_3)^{-1}E_t\left[\sum_{j=0}^{\infty}\lambda_3^{-j}\hat{f}_{t+j}\right] \\ &= (\beta\gamma\lambda_3)^{-1}E_t[(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_t], \end{aligned}$$

which corresponds to Equation (A10).

We now show that Equation (A10) implies Equation (A13). Since Equation (A10) holds for all $t \geq t_0$, Equation (A10) implies

$$(1 - \lambda_1 L)(1 - \lambda_2 L)\hat{i}_{t+1} = (\beta\gamma\lambda_3)^{-1}E_{t+1}[(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_{t+1}].$$

Multiplying by $\beta\gamma(1 - \lambda_3 L)$, and taking expectations at date t on both sides, and using Factorization (20), we obtain

$$\begin{aligned} E_t[A(L)\hat{i}_{t+1}] &= \lambda_3^{-1}E_t[(1 - \lambda_3 L)(1 - \lambda_3^{-1}L^{-1})^{-1}\hat{f}_{t+1}] \\ &= E_t[(1 - \lambda_3^{-1}L^{-1})(1 - \lambda_3^{-1}L^{-1})^{-1}L\hat{f}_{t+1}] \\ &= -E_t[L\hat{f}_{t+1}] = -\hat{f}_t \end{aligned}$$

which corresponds to Equation (A13).

Rule II. To show that a pair of bounded processes $\{\hat{i}_t, \hat{f}_t\}$ satisfies Equation (A11) at all dates if and only they satisfy Equation (A13) at all dates, we simply need to repeat the above steps, replacing λ_2 with λ_3 and vice versa.

Rule III. Let us now allow λ_2, λ_3 to be either real values, or complex conjugates, lying outside the unit circle. Since Equation (A13) implies both Equations (A10) and (A11), we know that Equation (A13) implies

$$(1 - \lambda_1 L) \left[\frac{(1 - \lambda_2 L) + (1 - \lambda_3 L)}{2} \right] \hat{i}_t = \frac{1}{2} (\beta \gamma \lambda_3)^{-1} E_t [(1 - \lambda_3^{-1} L^{-1})^{-1} \hat{f}_t] + \frac{1}{2} (\beta \gamma \lambda_2)^{-1} E_t [(1 - \lambda_2^{-1} L^{-1})^{-1} \hat{f}_t],$$

which is obtained by summing Equations (A10) and (A11) on both sides and dividing by 2. Thus, Equation (A13) implies Equation (A12).

We now show that Equation (A12) implies Equation (A13). Since Equation (A12) holds for all $t \geq t_0$, Equation (A12) implies

$$(1 - \lambda_1 L) \left(1 - \frac{\lambda_2 + \lambda_3}{2} L \right) \hat{i}_{t+2} = (2\beta\gamma\lambda_3)^{-1} E_{t+2} [(1 - \lambda_3^{-1} L^{-1})^{-1} \hat{f}_{t+2}] + (2\beta\gamma\lambda_2)^{-1} E_{t+2} [(1 - \lambda_2^{-1} L^{-1})^{-1} \hat{f}_{t+2}].$$

Multiplying by $\beta\gamma(1 - \lambda_2 L)(1 - \lambda_3 L)$, and taking expectations at date t on both sides, and using Factorization (20), we obtain

$$\begin{aligned} E_t \left[A(L) \left(1 - \frac{\lambda_2 + \lambda_3}{2} L \right) \hat{i}_{t+2} \right] &= \frac{1}{2\lambda_3} E_t [(1 - \lambda_2 L)(1 - \lambda_3 L)(1 - \lambda_3^{-1} L^{-1})^{-1} \hat{f}_{t+2}] \\ &\quad + \frac{1}{2\lambda_2} E_t [(1 - \lambda_2 L)(1 - \lambda_3 L)(1 - \lambda_2^{-1} L^{-1})^{-1} \hat{f}_{t+2}] \\ &= -\frac{1}{2} E_t [(1 - \lambda_2 L) \hat{f}_{t+1}] - \frac{1}{2} E_t [(1 - \lambda_3 L) \hat{f}_{t+1}] \\ &= -E_t \left[\left(1 - \frac{\lambda_2 + \lambda_3}{2} L \right) \hat{f}_{t+1} \right]. \end{aligned}$$

It follows that

$$-E_t \left[A(L) \left(1 - \frac{2}{\lambda_2 + \lambda_3} L^{-1} \right) L \hat{i}_{t+2} \right] = E_t \left[\left(1 - \frac{2}{\lambda_2 + \lambda_3} L^{-1} \right) L \hat{f}_{t+1} \right],$$

and hence that

$$E_t [A(L)(1 - \alpha L^{-1}) \hat{i}_{t+1}] = v_t,$$

where $0 \leq \alpha \equiv (2\lambda_2 + \lambda_3) < 1$ and $v_t \equiv -E_t [(1 - \alpha L^{-1}) \hat{f}_t]$. This implies furthermore

$$\begin{aligned} E_t [A(L) \hat{i}_{t+1}] &= \alpha E_t [A(L) \hat{i}_{t+2}] + v_t \\ &= E_t \left[\sum_{j=0}^{\infty} \alpha^j v_{t+j} \right] = E_t [(1 - \alpha L^{-1})^{-1} v_t] = -\hat{f}_t, \end{aligned}$$

which corresponds to Equation (A13).

A4. Proposition 4

PROPOSITION 4: Under the assumptions of Proposition 2, and a loss function with $\lambda_x, \lambda_i > 0$, each of the Rules I, II, and III has a representation of the form

$$i_t = (1 - \rho_1)i_t^* + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_\pi F_t(\pi) + \frac{\phi_x}{4} F_t(x) - \theta_\pi \pi_{t-1} - \frac{\theta_x}{4} x_{t-1} \tag{A14}$$

for all values of γ for which the rule is well defined, and in this representation

$$\rho_1 > 1, \rho_2 > 0,$$

$$0 < \theta_\pi \leq \phi_\pi,$$

and

$$0 < \theta_x = \phi_x.$$

Furthermore, for given values of the other parameters, as $\gamma \rightarrow 0$ (for Rule I) the coefficient θ_π approaches zero, though ϕ_π approaches a positive limit; while as $\gamma \rightarrow 1$ (for Rule III) the coefficients θ_π and ϕ_π approach the same limit.

PROOF: Proposition 2 guarantees that the roots $\lambda_1, \lambda_2, \lambda_3$ in the Factorization (20) are either real and satisfy $0 < \lambda_1 < 1 < \lambda_2 \leq \lambda_3$, or $0 < \lambda_1 < 1$, and λ_2, λ_3 are complex conjugates that lie outside the unit circle. Consider first the case in which $\lambda_1, \lambda_2, \lambda_3$ are real, so that both Rule I and Rule II are well defined.

Rule I. First note that the Rule I, i.e., Relation (22) can be rewritten as

$$\hat{i}_t = \rho_1 \hat{i}_{t-1} + \rho_2 \Delta \hat{i}_{t-1} + (\beta\gamma\lambda_3)^{-1} v_t, \tag{A15}$$

where

$$\rho_1 = 1 + (\lambda_2 - 1)(1 - \lambda_1) > 1,$$

$$\rho_2 = \lambda_1 \lambda_2 > 0,$$

and

$$v_t \equiv E_t[(1 - \lambda_3^{-1} L^{-1})^{-1} \hat{f}_t] = E_t \left[\sum_{j=0}^{\infty} \lambda_3^{-j} \hat{f}_{t+j} \right].$$

Since $E_t \hat{f}_{t+j}$ is given by

$$\begin{aligned} E_t \hat{f}_{t+j} &= \frac{\kappa\sigma}{\lambda_i} E_t (\hat{q}_{t+j} - \beta\gamma \hat{q}_{t+j+1}) \\ &= \frac{\kappa\sigma}{\lambda_i} E_t [-\gamma \hat{\pi}_{t+j-1} + (1 + \beta\gamma^2) \hat{\pi}_{t+j} - \beta\gamma \hat{\pi}_{t+j+1}] \\ &\quad + \frac{\lambda_x \sigma}{\lambda_i} E_t [-\hat{x}_{t+j-1} + (1 + \beta\gamma) \hat{x}_{t+j} - \beta\gamma \hat{x}_{t+j+1}], \end{aligned}$$

we have

$$\begin{aligned}
 v_t &= \frac{\kappa\sigma}{\lambda_i} E_t \left[\sum_{j=0}^{\infty} \lambda_3^{-j} (-\gamma \hat{\pi}_{t+j-1} + (1 + \beta\gamma^2) \hat{\pi}_{t+j} - \beta\gamma \hat{\pi}_{t+j+1}) \right] \\
 &\quad + \frac{\lambda_x\sigma}{\lambda_i} E_t \left[\sum_{j=0}^{\infty} \lambda_3^{-j} (-\hat{x}_{t+j-1} + (1 + \beta\gamma) \hat{x}_{t+j} - \beta\gamma \hat{x}_{t+j+1}) \right] \\
 &= \frac{\kappa\sigma}{\lambda_i} \sum_{j=-1}^{\infty} \tilde{\alpha}_{\pi,j} E_t \hat{\pi}_{t+j} + \frac{\lambda_x\sigma}{\lambda_i} \sum_{j=-1}^{\infty} \alpha_{x,j} E_t \hat{x}_{t+j},
 \end{aligned}$$

where

$$\tilde{\alpha}_{\pi,-1} = -\gamma, \tag{A16}$$

$$\tilde{\alpha}_{\pi,0} = 1 + \beta\gamma^2 - \lambda_3^{-1}\gamma, \tag{A17}$$

$$\tilde{\alpha}_{\pi,j} = -\lambda_3^{-j+1}\beta\gamma + \lambda_3^{-j}(1 + \beta\gamma^2) - \lambda_3^{-j-1}\gamma, \forall j \geq 1, \tag{A18}$$

and

$$\alpha_{x,-1} = -1, \tag{A19}$$

$$\alpha_{x,0} = 1 + \beta\gamma - \lambda_3^{-1}, \tag{A20}$$

$$\alpha_{x,j} = -\lambda_3^{-j+1}\beta\gamma + \lambda_3^{-j}(1 + \beta\gamma) - \lambda_3^{-j-1}, \forall j \geq 1. \tag{A21}$$

The variable v_t can furthermore be written as

$$v_t = \frac{\kappa\sigma}{\lambda_i} S_\pi \sum_{j=0}^{\infty} \alpha_{\pi,j} E_t \hat{\pi}_{t+j} + \frac{\lambda_x\sigma}{\lambda_i} \sum_{j=0}^{\infty} \alpha_{x,j} E_t \hat{x}_{t+j} - \frac{\kappa\sigma\gamma}{\lambda_i} \hat{\pi}_{t-1} - \frac{\lambda_x\sigma}{\lambda_i} \hat{x}_{t-1}, \tag{A22}$$

where

$$\begin{aligned}
 S_\pi &= \sum_{j=0}^{\infty} \tilde{\alpha}_{\pi,j} \\
 &= -(0 + \lambda_3^{-0} + \lambda_3^{-1} + \lambda_3^{-2} + \dots)\beta\gamma + (\lambda_3^{-0} + \lambda_3^{-1} + \lambda_3^{-2} + \dots)(1 + \beta\gamma^2) \\
 &\quad - (\lambda_3^{-1} + \lambda_3^{-2} + \dots)\gamma = (1 - \lambda_3^{-1})^{-1}(1 + \beta\gamma^2 - \beta\gamma - \lambda_3^{-1}\gamma)
 \end{aligned}$$

and

$$\alpha_{\pi,j} = \frac{\tilde{\alpha}_{\pi,j}}{S_\pi}, \forall j \geq 0.$$

Note that the coefficients $\alpha_{\pi,j}$ satisfy

$$\sum_{j=0}^{\infty} \alpha_{\pi,j} = S_\pi^{-1} \sum_{j=0}^{\infty} \tilde{\alpha}_{\pi,j} = 1,$$

and the coefficients $\alpha_{x,j}$ satisfy

$$\sum_{j=0}^{\infty} \alpha_{xj} = -(\lambda_3^{-0} + \lambda_3^{-1} + \lambda_3^{-2} + \dots)\beta\gamma + (\lambda_3^{-0} + \lambda_3^{-1} + \lambda_3^{-2} + \dots)(1 + \beta\gamma) - (\lambda_3^{-1} + \lambda_3^{-2} + \dots) = (1 - \lambda_3^{-1})^{-1}(1 - \lambda_3^{-1}) = 1 .$$

Combining Equation (A15) and Equation (A22), we can rewrite Rule I as

$$\hat{i}_t = \rho_1 \hat{i}_{t-1} + \rho_2 \Delta \hat{i}_{t-1} + \phi_\pi F_t(\hat{\pi}) + \frac{\phi_x}{4} F_t(\hat{x}) - \theta_\pi \hat{\pi}_{t-1} - \frac{\theta_x}{4} \hat{x}_{t-1} , \tag{A23}$$

where

$$\phi_\pi = (\beta\gamma\lambda_3)^{-1} \frac{\kappa\sigma}{\lambda_i} S_\pi = \frac{\kappa\sigma}{\lambda_i\beta} \frac{1 + \beta\gamma^2 - \beta\gamma - \lambda_3^{-1}\gamma}{\lambda_3\gamma(1 - \lambda_3^{-1})} ,$$

$$\theta_\pi = (\beta\gamma\lambda_3)^{-1} \frac{\kappa\sigma\gamma}{\lambda_i} = \frac{\kappa\sigma}{\lambda_i\beta\lambda_3} > 0 ,$$

and

$$\phi_x = \theta_x = \frac{4\lambda_x\sigma}{\lambda_i\beta\gamma\lambda_3} > 0 .$$

Note furthermore that

$$\phi_\pi = \theta_\pi \frac{\lambda_3 + \beta\gamma^2\lambda_3 - \beta\gamma\lambda_3 - \gamma}{\gamma(\lambda_3 - 1)} = \theta_\pi \left(1 + \frac{(1 - \gamma)(1 - \beta\gamma)}{\gamma(1 - \lambda_3^{-1})} \right) \geq \theta_\pi . \tag{A24}$$

Recalling that $\hat{z}_t \equiv z_t - \bar{z}$ for any variable z , and that $\phi_x = \theta_x$, we can rewrite Equation (A23) as

$$i_t = (1 - \rho_1)\bar{i} - (\phi_\pi - \theta_\pi)\bar{\pi} + \rho_1 i_{t-1} + \rho_2 \Delta i_{t-1} + \phi_\pi F_t(\pi) + \frac{\phi_x}{4} F_t(x) - \theta_\pi \pi_{t-1} - \frac{\theta_x}{4} x_{t-1} . \tag{A25}$$

We know from Proposition 3 that Equation (21) holds, and thus that Equation (16) holds. In the steady state, Equation (16) reduces to

$$A(L)(\bar{i} - i^*) = -\bar{f} ,$$

where

$$\bar{f} = \frac{\kappa\sigma}{\lambda_i}(1 - \beta\gamma)\bar{q} = \frac{\kappa\sigma}{\lambda_i}(1 - \beta\gamma)(1 - \gamma)\bar{\pi} .$$

It follows from Factorization (20) that

$$(1 - \lambda_1)(1 - \lambda_2)i^* = (1 - \lambda_1)(1 - \lambda_2)\bar{i} + \frac{\kappa\sigma}{\lambda_i} \frac{(1 - \gamma)(1 - \beta\gamma)}{\beta\gamma(1 - \lambda_3)}\bar{\pi} . \tag{A26}$$

Given that

$$(1 - \lambda_1)(1 - \lambda_2) = 1 - \rho_1$$

and given that Equation (A24) implies

$$\frac{\kappa\sigma}{\lambda_i} \frac{(1 - \gamma)(1 - \beta\gamma)}{\beta\gamma(1 - \lambda_3)} = -(\phi_\pi - \theta_\pi),$$

we can rewrite Equation (A26) as

$$(1 - \rho_1)i^* = (1 - \rho_1)\bar{i} - (\phi_\pi - \theta_\pi)\bar{\pi}.$$

Combining this with Equation (A25) yields Equation (A14).

As γ approaches 0, we have $\lambda_3^{-1} \rightarrow 0$ and $\lambda_3\gamma \rightarrow \beta^{-1}$. It follows that

$$\lim_{\gamma \rightarrow 0} \phi_\pi = \lim_{\gamma \rightarrow 0} \frac{\kappa\sigma}{\lambda_i\beta} \frac{1 + \beta\gamma^2 - \beta\gamma - \lambda_3^{-1}\gamma}{\lambda_3\gamma(1 - \lambda_3^{-1})} = \frac{\kappa\sigma}{\lambda_i} > 0,$$

$$\lim_{\gamma \rightarrow 0} \theta_\pi = \lim_{\gamma \rightarrow 0} \frac{\kappa\sigma}{\lambda_i\beta} \lambda_3^{-1} = 0,$$

and

$$\lim_{\gamma \rightarrow 0} \phi_x = \lim_{\gamma \rightarrow 0} \theta_x = \lim_{\gamma \rightarrow 0} \frac{4\lambda_x\sigma}{\lambda_i\beta\gamma\lambda_3} = \frac{4\lambda_x\sigma}{\lambda_i} > 0.$$

Rule II. Following the same development as for Rule I, but replacing λ_2 with λ_3 and vice versa, we can show that Relation (23) can also be written as in Equation (A14), but where

$$\rho_1 = 1 + (\lambda_3 - 1)(1 - \lambda_1) > 1,$$

$$\rho_2 = \lambda_1\lambda_3 > 0,$$

$$\theta_\pi = \frac{\kappa\sigma}{\lambda_i\beta\lambda_2} > 0,$$

$$\phi_\pi = \theta_\pi \left(1 + \frac{(1 - \gamma)(1 - \beta\gamma)}{\gamma(1 - \lambda_2^{-1})} \right) \geq \theta_\pi,$$

and

$$\phi_x = \theta_x = \frac{4\lambda_x\sigma}{\lambda_i\beta\gamma\lambda_2} > 0.$$

Rule III. We now allow the roots λ_2 and λ_3 to be either real or complex. Recall from the proof of Proposition 2 that $(\lambda_2 + \lambda_3)/2$ is real and is greater than 1. Using this, Rule III (Equation 24) can be rewritten, for all values of $\gamma \in (0, 1]$, as

$$\hat{t}_t = \rho_1 \hat{t}_{t-1} + \rho_2 \Delta \hat{t}_{t-1} + \frac{1}{2} (\beta \gamma \lambda_3)^{-1} v_t^I + \frac{1}{2} (\beta \gamma \lambda_2)^{-1} v_t^{II}, \tag{A27}$$

where ρ_1 and ρ_2 are now given by

$$\rho_1 = \lambda_1 + \frac{\lambda_2 + \lambda_3}{2} - \lambda_1 \frac{\lambda_2 + \lambda_3}{2} = 1 + \left(\frac{\lambda_2 + \lambda_3}{2} - 1 \right) (1 - \lambda_1) > 1, \tag{A28}$$

$$\rho_2 = \lambda_1 \frac{\lambda_2 + \lambda_3}{2} > 0, \tag{A29}$$

and

$$\begin{aligned} v_t^I &\equiv E_t[(1 - \lambda_3^{-1} L^{-1})^{-1} \hat{f}_t] \\ &= \frac{\kappa \sigma}{\lambda_i} \sum_{j=-1}^{\infty} \tilde{\alpha}_{\pi,j}^I E_t \hat{\pi}_{t+j} + \frac{\lambda_x \sigma}{\lambda_i} \sum_{j=-1}^{\infty} \alpha_{x,j}^I E_t \hat{x}_{t+j}, \end{aligned} \tag{A30}$$

$$\begin{aligned} v_t^{II} &\equiv E_t[(1 - \lambda_2^{-1} L^{-1})^{-1} \hat{f}_t] \\ &= \frac{\kappa \sigma}{\lambda_i} \sum_{j=-1}^{\infty} \tilde{\alpha}_{\pi,j}^{II} E_t \hat{\pi}_{t+j} + \frac{\lambda_x \sigma}{\lambda_i} \sum_{j=-1}^{\infty} \tilde{\alpha}_{x,j}^{II} E_t \hat{x}_{t+j}, \end{aligned} \tag{A31}$$

and where $\tilde{\alpha}_{\pi,j}^I$ and $\alpha_{x,j}^I$ are defined in Equations (A16)–(A21) for all $j \geq 1$, $\tilde{\alpha}_{\pi,j}^{II}$ and $\alpha_{x,j}^{II}$ are defined in the same way except that λ_3 is replaced with λ_2 . Using Equations (A30) and (A31), Equation (A27) can furthermore be written as

$$\begin{aligned} \hat{t}_t &= \rho_1 \hat{t}_{t-1} + \rho_2 \Delta \hat{t}_{t-1} + \frac{\kappa \sigma}{\lambda_i \beta \gamma} \frac{S_\pi}{2} \sum_{j=0}^{\infty} \alpha_{\pi,j} E_t \hat{\pi}_{t+j} \\ &+ \frac{\lambda_x \sigma}{\lambda_i \beta \gamma} \frac{S_x}{2} \sum_{j=0}^{\infty} \alpha_{x,j} E_t \hat{x}_{t+j} - \frac{\kappa \sigma}{\lambda_i \beta} \frac{\lambda_2^{-1} + \lambda_3^{-1}}{2} \hat{\pi}_{t-1} - \frac{\lambda_x \sigma}{\lambda_i \beta \gamma} \frac{\lambda_2^{-1} + \lambda_3^{-1}}{2} \hat{x}_{t-1}, \end{aligned} \tag{A32}$$

where

$$\begin{aligned} S_\pi &= \sum_{j=0}^{\infty} (\lambda_3^{-1} \tilde{\alpha}_{\pi,j}^I + \lambda_2^{-1} \tilde{\alpha}_{\pi,j}^{II}), \\ S_x &= \sum_{j=0}^{\infty} (\lambda_3^{-1} \alpha_{x,j}^I + \lambda_2^{-1} \alpha_{x,j}^{II}) = \lambda_2^{-1} + \lambda_3^{-1}, \\ \alpha_{\pi,j} &= S_\pi^{-1} (\lambda_3^{-1} \tilde{\alpha}_{\pi,j}^I + \lambda_2^{-1} \tilde{\alpha}_{\pi,j}^{II}), \\ \alpha_{x,j} &= S_x^{-1} (\lambda_3^{-1} \alpha_{x,j}^I + \lambda_2^{-1} \alpha_{x,j}^{II}), \end{aligned}$$

and where

$$\sum_{j=0}^{\infty} \alpha_{\pi,j} = \sum_{j=0}^{\infty} \alpha_{x,j} = 1.$$

Equation (A32) is of the Form (A23), where ρ_1 and ρ_2 are defined in Equations (A28), (A29), and

$$\phi_\pi = \frac{\kappa\sigma}{\lambda_i\beta} \frac{S_\pi}{2}, \tag{A33}$$

$$\theta_\pi = \frac{\kappa\sigma}{\lambda_i\beta} \frac{\lambda_2^{-1} + \lambda_3^{-1}}{2} > 0, \tag{A34}$$

$$\phi_x = \theta_x = \frac{4\lambda_x\sigma}{\lambda_i\beta\gamma} \frac{\lambda_2^{-1} + \lambda_3^{-1}}{2} > 0. \tag{A35}$$

(Note that if λ_2 and λ_3 are complex conjugates, then $\lambda_2^{-1} + \lambda_3^{-1} = (\lambda_2 + \lambda_3)/(\lambda_2\lambda_3)$ is real.) Note furthermore that for all $\gamma \in (0, 1]$, the coefficient ϕ_π satisfies

$$\begin{aligned} \phi_\pi &= \frac{\kappa\sigma}{\lambda_i\beta} \left(\frac{\lambda_3^{-1}}{2} \frac{1 + \beta\gamma^2 - \beta\gamma - \lambda_3^{-1}\gamma}{\gamma(1 - \lambda_3^{-1})} + \frac{\lambda_2^{-1}}{2} \frac{1 + \beta\gamma^2 - \beta\gamma - \lambda_2^{-1}\gamma}{\gamma(1 - \lambda_2^{-1})} \right) \\ &= \frac{\kappa\sigma}{\lambda_i\beta} \left(\frac{\lambda_3^{-1}}{2} \left(1 + \frac{(1 - \gamma)(1 - \beta\gamma)}{\gamma(1 - \lambda_3^{-1})} \right) + \frac{\lambda_2^{-1}}{2} \left(1 + \frac{(1 - \gamma)(1 - \beta\gamma)}{\gamma(1 - \lambda_2^{-1})} \right) \right) \\ &= \theta_\pi - \frac{\kappa\sigma}{\lambda_i\beta\gamma} \frac{(1 - \gamma)(1 - \beta\gamma)}{(1 - \lambda_2)(1 - \lambda_3)} \left(1 - \frac{\lambda_2 + \lambda_3}{2} \right) \geq \theta_\pi. \end{aligned} \tag{A36}$$

As for Rule I, we can rewrite Equation (A23) as in Equation (A25), but where the coefficients are given in Equations (A28), (A29), and (A33)–(A35). Again, we know from Proposition 3 that Equation (16) holds, and thus that Equation (A26) holds. Multiplying Equation (A26) on both sides by $(1 - \lambda_2)^{-1}(1 - (\lambda_2 + \lambda_3)/2)$, we obtain

$$\begin{aligned} (1 - \lambda_1) \left(1 - \frac{\lambda_2 + \lambda_3}{2} \right) i^* &= (1 - \lambda_1) \left(1 - \frac{\lambda_2 + \lambda_3}{2} \right) \bar{i} \\ &+ \frac{\kappa\sigma}{\lambda_i\beta\gamma} \frac{(1 - \gamma)(1 - \beta\gamma)}{(1 - \lambda_2)(1 - \lambda_3)} \left(1 - \frac{\lambda_2 + \lambda_3}{2} \right) \bar{\pi}. \end{aligned} \tag{A37}$$

Since Equations (A28) and (A36) imply

$$\begin{aligned} (1 - \lambda_1) \left(1 - \frac{\lambda_2 + \lambda_3}{2} \right) &= 1 - \rho_1, \\ \frac{\kappa\sigma}{\lambda_i\beta\gamma} \frac{(1 - \gamma)(1 - \beta\gamma)}{(1 - \lambda_2)(1 - \lambda_3)} \left(1 - \frac{\lambda_2 + \lambda_3}{2} \right) &= -(\phi_\pi - \theta_\pi), \end{aligned}$$

we can rewrite Equation (A37) as

$$(1 - \rho_1) i^* = (1 - \rho_1) \bar{i} - (\phi_\pi - \theta_\pi) \bar{\pi}.$$

Combining this with Equation (A25) yields Equation (A14).

Finally, in the limit, as $\gamma = 1$, we have

$$\phi_{\pi} = \frac{\kappa\sigma}{\lambda_i\beta} \frac{\lambda_2^{-1} + \lambda_3^{-1}}{2} = \theta_{\pi}.$$

NOTES

1. See, e.g., Coletti et al. (1996) and Black et al. (1997).
2. See, e.g., Vickers (1998).
3. Examples of studies of this kind that analyze forecast-based policy rules include Batini and Haldane (1999), Levin, Wieland, and Williams (2003), and Batini and Pearlman (2002).
4. This includes our own previous studies, such as Rotemberg and Woodford (1999).
5. In the context of the model of Section 2 below, *any* specification of the path of the central bank's nominal interest-rate instrument as a function solely of the history of exogenous disturbances implies indeterminacy of equilibrium, as shown in Woodford (2003, chapter 4). This problem with exogenous specifications of an interest-rate path is common to many optimizing models of the monetary transmission mechanism, and occurs for essentially the same reason as in the classic analysis of Sargent and Wallace (1975).
6. While the general approach to the construction of robustly optimal policy rules used here is the same as that discussed in the companion paper, the derivations presented here are self-contained and do not rely upon any of the results for the general linear-quadratic problem presented in the earlier paper. It is our hope that a self-contained exposition of the relevant calculations for these simple models will serve to increase insight into the method, in addition to delivering results of interest with regard to these particular models.
7. The microeconomic foundations of the structural relations assumed here are expounded in Woodford (2003, chapters 3 and 4).
8. See Woodford (2003, chapter 6) for discussion of the welfare-relevant output gap and of the nature of "cost-push shocks."
9. The assumption here of a concern for interest-rate stabilization deserves particular comment. This may be justified in either of two ways. First, transaction frictions of the sort that account for a demand for the monetary base despite the fact that it earns no interest result in a welfare loss that is (to second order) proportional to the squared differential between the interest available on other assets and that earned on the monetary base; in this case, $i^* = 0$ and λ_i can be calibrated based on an estimated money demand function. Alternatively, a quadratic penalty on interest-rate variations can approximate (within a linear-quadratic framework) the effects of the requirement that nominal interest rates never be negative; in this case, i^* may be slightly higher than the rate of time preference (the nominal interest rate consistent with a zero-inflation steady state), and the size of λ_i depends on the frequency and extent to which the natural rate of interest is sometimes negative. In Woodford (1999) and below, λ_i is calibrated on the latter ground, using the natural-rate process estimated in Rotemberg and Woodford (1999).
10. In terms of the notation of Giannoni and Woodford (2002, section 5), we assume here that $t_0 = 0$.
11. This is a substantial advantage of this instrument rule over the one proposed in Woodford (1999), which expresses the federal funds rate as a function of the lagged funds rate, the lagged rate of increase in the funds rate, the current inflation rate, and the previous quarter's inflation rate. This rule would also be consistent with optimal responses to real disturbances, but only if (as assumed in the earlier calculation) all disturbances perturb the natural rate of interest in a way that can be described by an AR(1) process with a single specified coefficient of serial correlation, and have no effect on the natural rate of output that is different than the effect on the efficient rate of output (i.e., there are no cost-push shocks). In this special case, however, the rule discussed earlier has the advantage that its implementation requires no information on the part of the central bank other than an accurate measure of inflation (including an accurate projection of period t inflation at the time that the period t funds rate is set).
12. See the discussion in Woodford (2003, chapter 4, section 2.2) of the generalization of this principle to the case of policy rules with interest-rate inertia.
13. It might appear that the history dependence of optimal policy here depends critically on our assumption of a loss function with $\lambda_i > 0$. But Eggertsson and Woodford (2003) consider the optimal nonlinear policy rule for a cashless economy, from the point of view of an objective that assigns no penalty to interest-rate variations, and instead impose the zero bound on nominal interest rates as a nonlinear constraint each period. They again find that the optimal policy rule is history dependent; acceptable inflation and output-gap projections depend on whether the zero bound has been a binding constraint in previous quarters. Aoki (2002) similarly assigns no penalty to interest-rate variations, but assumes that the central bank has incomplete information about the current state of the economy when

setting its interest-rate operating target. Here again, the optimal policy rule is history dependent; optimal policy involves a commitment to respond later when the central bank learns of errors in its past estimates, even with regard to state variables that no longer matter for inflation or output determination. As long as there is *some* reason why the central bank cannot completely achieve its stabilization objectives independently in each period, optimal policy is history dependent.

14. It should also be noted that the output-gap measure used in Judd and Rudebusch's empirical analysis, while a plausible measure of what the Fed is likely to have responded to, may not correspond to the welfare-relevant output gap indicated by the variable x_t in the optimal Rule (9). In addition, ϕ_π indicates response to the most recent four-quarter growth in the GDP deflator, rather than an annualized inflation rate over the past quarter alone.

15. Here it may be noted that Rule (9) is not a *uniquely* optimal policy rule; it is not even the only rule that is robustly optimal in the sense that we discuss here. In Giannoni and Woodford (2004, section 1.3), we discuss an alternative rule, a *pure targeting rule* (i.e., a criterion that the projected paths of inflation and the output gap must fulfill at all times, that does not involve the current instrument setting) that is equivalent to Rule (9) in the sense that nonexplosive paths for inflation, the output gap, and the nominal interest rate will be projected to satisfy the target criterion at all times if and only if they are projected to satisfy Rule (9) at all times. This alternative representation of optimal policy, unlike Rule (9), involves forecasts of inflation and the output gap in quarters beyond the one for which policy is currently being chosen. However, even in this case, the relevant forecasts do not look too far into the future; the greatest weight is placed on the projection for the current quarter, and in the case of the parameter values given in Table 1, the mean forecast horizon is only 2.1 quarters in the future. Perhaps more to the point, it is not *necessary* to represent the policy rule in even such a forward-looking form as that, in order to have a robustly optimal policy rule; for Rule (9) is one possible representation of the optimal rule.

16. This is obviously not the case if, as more realistic models often assume, there are delays in the effect of any new information on prices and spending. But in this case, it is probably not desirable for overnight interest rates to respond immediately to news, either; see Section 4 below.

17. This is not true if there are delays in the effects of shocks upon inflation and output, as discussed in Section 4 below. But in that case, even the *delayed* effect upon the central bank's instrument that is required by optimal policy cannot be implemented on the basis only of lagged observations of the target variables, because of the delay with which shocks affect these variables.

18. Christiano et al. assume complete indexation of prices to the previous quarter's aggregate price index, while we allow for partial indexation, measured by a coefficient γ , following Smets and Wouters (2003). Christiano et al. also assume that wages as well as prices are sticky, and that wages are also indexed to the lagged price index, as in the model discussed in Section 3. Here we continue, as in the previous section, to assume flexible wages, some other form of efficient contracting in the labor market, or direct supply of output by "yeoman farmers." The same kind of indexation is also assumed in Altig et al. (2002).

19. An alternative way of modeling inflation inertia would be to assume the existence of backward-looking "rule of thumb" price setters, as in Gali and Gertler (1999). This leads to a modification of the aggregate supply relation that is similar, though not quite identical, to Equation (12). Steinsson (2000) and Amato and Laubach (2001) derive welfare-theoretic loss functions for this model, and find that the loss each period is a quadratic function of both π_t and π_{t-1} that is similar, though again not identical, to our Loss function (13). Hence, we conjecture that similar conclusions as to the degree to which optimal policy is forward looking would be obtained using the Gali-Gertler model, though we do not take this up here.

20. A rule expressed in this way will also conform better to the evident preference of central banks to justify their monetary policy decisions to the public in terms of their projections for the future paths of inflation and output, rather than in terms of their assumptions about the future path of interest rates. Public communications such as the Bank of England's *Inflation Report* put projections for both inflation and output at center stage, while being careful not to express any opinion whatsoever about the likely path of interest rates over the period under discussion. The forecast-based rules proposed below still refer to forecast paths conditional upon intended policy, rather than upon "constant-interest-rate" forecasts, and so it will not be possible to implement these rules without taking a stand (at least for internal purposes) on the likely future path of interest rates. But the rules make it possible to discuss the way in which the current instrument setting is required by the bank's inflation and output projections, without also discussing the interest-rate path that is implicit in those projections, and to this extent they require a less radical modification of current procedures.

21. This proposed selection principle also chooses the rule that results in the least intrinsic inertia in the implied interest-rate dynamics; this selection is appropriate if we wish to establish the extent to which the optimal interest-rate rule is *necessarily* inertial.

22. Note that as $\gamma \rightarrow 0$, $\lambda_3 \rightarrow +\infty$, while $\gamma\lambda_3 \rightarrow \beta^{-1}$.

23. Here we plot the relative weights, rather than the absolute weights, because this makes visual comparison between the degree of forecast dependence of optimal policy in the different cases easier. The absolute weights can be recovered by integrating the plots shown here, since the relative weights in each case must sum to $1/\alpha_{z,0}$.

24. The loss function assumed here is derived in Woodford (2003, chapter 6). Our loss function differs slightly from the one derived by Erceg et al., even for the case that they treat, because they do not take account of the discounting of utility in the way that we do.

25. Here we have rewritten $\log Y_t - \log Y_t^a$ as $x_t + u_t$, where as in our basic model, x_t is the gap between actual and *efficient* output and u_t represents inefficient variation in the natural rate of output.

26. The welfare-theoretic loss function for this model when $\gamma_w = \gamma_p = 0$ and there is no penalty for interest-rate variability is derived in Woodford (2003, chapter 6, section 4.4). When $\gamma_w, \gamma_p > 0$, the relation of wage and price dispersion to wage and price inflation changes in the way discussed in Woodford (2003, chapter 6, section 2.2), resulting in the modification indicated here of the first two terms of the loss function. The justification for the final term is the same as in the flexible-wage model above.

27. Note that if $\kappa_w = \kappa_p$, as assumed here, then the real wage is unaffected by monetary policy, as discussed in Woodford (2003, chapter 3, section 4.2). In this case the Rotemberg–Woodford inflation equation is correctly specified even when wages are sticky (though their welfare analysis would not be correct), and their parameter “ κ ” corresponds to κ_p here.

28. Note that Sbordone’s inflation equation is equally valid regardless of whether wages are sticky or not.

29. Even though our parameter values have been taken from two different studies using different data sets, the implied values of these parameters are reasonable and not too different from the estimates of Amato and Laubach (2003); see, for example, table 4.2 of Woodford (2003).

30. If λ_i resulted solely from the existence of transactions frictions, as discussed in Woodford (2003, chapter 6, section 4.1), the same calibrated value would be appropriate regardless of the assumed degree of wage stickiness. In the case that λ_i is chosen to reflect the advantages of lower interest-rate variability as a result of the zero bound, as in Woodford (1999), then the appropriate value would depend on the assumed variance of disturbances. In this case, the appropriate value is not independent of whether we assume wages to be sticky, because the other stabilization objectives are not the same in this case; but we do not here consider the degree to which the appropriate value of λ_i should change.

31. This explains why we plot relative weights rather than the weights $\alpha_{w,j}$ in Figure 4. If we normalize the $\alpha_{w,j}$ to sum to one, then the weights are undefined in the limiting case $\gamma = 1$. Nonetheless, the relative weights have well-defined limiting values, shown in the figure. The coefficients multiplying any given forecast of wage inflation—i.e., the products $\phi_w \alpha_{w,j}$ —also remain well defined, so there is a well-defined optimal policy rule in this case.

32. Alternatively, in Equations (36) and (37) we may interpret x_t to mean $\hat{Y}_t - E_{t-d} \hat{Y}_t^e$. In this case, the loss function given by Equations (3) and (13) is still correct, up to terms (involving the component of \hat{Y}_t^e that is not forecastable d periods in advance) that are independent of policy.

33. An IS relation of this form is presented in Woodford (2003, chapter 4, section 4.1), where the additional disturbance results from the unforecastable components of government purchases and/or the efficient level of output. In the case of inflation, one might suppose that wholesale prices are determined d periods in advance, and satisfy Equation (37), while the retail price of each good is equal to the wholesale price plus an exogenous markup, which markup need not be forecastable in advance.

LITERATURE CITED

- Altig, David, Lawrence J. Christiano, Martin S. Eichenbaum, and Jesper Linde (2002). “Technology Shocks and Aggregate Fluctuations.” Unpublished, Federal Reserve Bank of Cleveland.
- Amato, Jeffery D., and Thomas Laubach (2001). “Rule-of-Thumb Behavior and Monetary Policy.” Unpublished, Bank for International Settlements.
- Amato, Jeffery D., and Thomas Laubach (2003). “Estimation and Control of an Optimization-Based Model with Sticky Wages and Prices.” *Journal of Economic Dynamics and Control* 27, 1181–1215.
- Aoki, Kosuke (2002). “Optimal Commitment Policy under Noisy Information.” CEPR Discussion Paper No. 3370.
- Batini, Nicoletta, and Andrew G. Haldane (1999). “Forward-Looking Rules for Monetary Policy.” In *Monetary Policy Rules*, edited by J.B. Taylor. Chicago: The University of Chicago Press.
- Batini, Nicoletta, and Joe Pearlman (2002). “Too Much Too Soon: Instability and Indeterminacy with Forward-Looking Rules.” Unpublished, Bank of England.

- Black, Richard, Vincenzo Cassino, Aaron Drew, Eric Hansen, Benjamin Hunt, David Rose, and Alasdair Scott (1997). "The Forecasting and Policy System: The Core Model." Research Paper No. 43, Reserve Bank of New Zealand.
- Boivin, Jean, and Marc Giannoni (2003). "Has Monetary Policy Become Less Powerful?" NBER Working Paper No. 9459.
- Christiano, Lawrence J., Martin S. Eichenbaum, and Charles L. Evans (2001). "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." NBER Working Paper No. 8403.
- Clarida, Richard, Jordi Gali, and Mark Gertler (2000). "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *Quarterly Journal of Economics* 115, 147–180.
- Coletti, Donald, Benjamin Hunt, David Rose, and Robert Tetlow (1996). "The Dynamic Model: QPM, The Bank of Canada's New Quarterly Projection Model, Part 3," Technical Report No. 75, Bank of Canada.
- Eggertsson, Gauti B., and Michael Woodford (2003). "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity* 1, 139–211.
- Erceg, Christopher J., Dale W. Henderson, and Andrew T. Levin (2000). "Optimal Monetary Policy with Staggered Wage and Price Contracts." *Journal of Monetary Economics* 46, 281–313.
- Gali, Jordi, and Mark Gertler (1999). "Inflation Dynamics: A Structural Econometric Analysis." *Journal of Monetary Economics* 44, 195–222.
- Giannoni, Marc P. (2001). "Optimal Interest-Rate Rules in a Forward-Looking Model, and Inflation Stabilization versus Price-Level Stabilization." Unpublished, Princeton University.
- Giannoni, Marc P., and Michael Woodford (2002). "Optimal Interest-Rate Rules: I. General Theory." NBER Working Paper No. 9419.
- Giannoni, Marc P., and Michael Woodford (2004). "Optimal Inflation Targeting Rules." Forthcoming in *Inflation Targeting*, edited by Ben S. Bernanke and Michael Woodford. Chicago: The University of Chicago Press.
- Judd, John F., and Glenn D. Rudebusch (1998). "Taylor's Rule and the Fed: 1970–1997." *Federal Reserve Bank of San Francisco Economic Review* 3, 3–16.
- Levin, Andrew, Volker Wieland, and John C. Williams (2003). "The Performance of Forecast-Based Monetary Policy Rules Under Model Uncertainty." *American Economic Review* 93, 622–645.
- McCallum, Bennett T., and Edward Nelson (1999). "Performance of Operational Policy Rules in an Estimated Semi-classical Structural Model." In *Monetary Policy Rules*, edited by J.B. Taylor. Chicago: The University of Chicago Press.
- Nelson, Edward (2001). "U.K. Monetary Policy 1972–1997: A Guide Using Taylor Rules." Unpublished, Bank of England.
- Rotemberg, Julio J., and Michael Woodford (1997). "An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy." *NBER Macroeconomics Annual* 12, 297–346.
- Rotemberg, Julio J., and Michael Woodford (1999). "Interest-Rate Rules in an Estimated Sticky-Price Model." In *Monetary Policy Rules*, edited by J.B. Taylor. Chicago: The University of Chicago Press.
- Sbordone, Argia M. (2002). "Prices and Unit Labor Costs: A New Test of Price Stickiness." *Journal of Monetary Economics* 49, 265–292.
- Sargent, Thomas J., and Neil Wallace (1975). "'Rational' Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule," *Journal of Political Economy* 83, 241–254.

- Smets, Frank, and Raf Wouters (2003). "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association* 1, 1123–1175.
- Steinsson, Jon (2000). "Optimal Monetary Policy in an Economy with Inflation Persistence." Working Paper No. 11, Central Bank of Iceland.
- Svensson, Lars E.O. (1997). "Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets." *European Economic Review* 41, 1111–1146.
- Svensson, Lars E.O., and Michael Woodford (2004). "Implementing Optimal Policy through Inflation-Forecast Targeting." Forthcoming in *Inflation Targeting*, edited by Ben S. Bernanke and Michael Woodford. Chicago: The University of Chicago Press.
- Taylor, John B. (1993). "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39, 195–214.
- Vickers, John (1998). "Inflation Targeting in Practice: The U.K. Experience." *Bank of England Quarterly Bulletin* 38, 368–375.
- Woodford, Michael (1999). "Optimal Monetary Policy Inertia." NBER Working Paper No. 7261.
- Woodford, Michael (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.