Estimating Central Bank Preferences under Commitment and Discretion

Gregory E. Givens
Assistant Professor of Economics
Department of Economics and Finance
Jennings A. Jones College of Business
Middle Tennessee State University
Murfreesboro, TN 37132
ggivens@mtsu.edu

First Draft: June 2009
This Draft: October 2010

Abstract

This paper explains US macroeconomic outcomes with an empirical new-Keynesian model in which monetary policy minimizes the central bank's loss function. The presence of expectations in the model forms a well-known distinction between two modes of optimization, termed commitment and discretion. The model is estimated separately under each policy using maximum likelihood over the Volcker-Greenspan-Bernanke period. Comparisons of fit reveal that the data favor the specification with discretionary policy. Estimates of the loss function weights point to an excessive concern for interest rate smoothing in the commitment model but a more balanced concern relative to inflation and output stability in the discretionary model.

Keywords: Optimal Monetary Policy, Commitment, Discretion, Policy Preferences

JEL Classification: E52, E58, E61, C32, C61

1 Introduction

In forward-looking models that embody rational expectations, optimal monetary policies are separated by a dichotomy known as commitment and discretion. As first pointed out by Kydland and Prescott (1977), the key distinction between them is whether promises made at an earlier time restrict the policy choices of today. A central bank operating under commitment gives assurances about how policy will be set in all future periods through the design of an optimal contingency rule for the nominal interest rate. By a contingency rule I mean one involving instrument settings that are conditional on the state of the economy. The salient aspect of commitment is that policymakers deliver on past promises by always responding to economic conditions in accordance with the original plan. A central bank practicing discretion, however, is not bound by some predetermined course of action. Changes in the interest rate are instead the result of period-by-period reoptimization in which foregoing policy intentions are considered irrelevant for current decision making. It follows that measures taken to stabilize the economy do not constrain future policy management in any credible way.

Ever since the distinction between commitment and discretion was first recognized, countless studies have assessed their performance in a variety of macroeconomic models. A consistent theme of this literature is that commitment is the better policy because it generates lower average inflation in the long run (e.g., Barro and Gordon, 1983; Rogoff, 1985) and a more efficient response of the economy to random shocks in the short run (e.g., Woodford, 1999; Clarida, Galí, and Gertler, 1999). The gains from commitment are a direct consequence of the role that expectations play in shaping economic conditions. A policymaker that is understood by private agents to always follow through on promised behavior can harness expectations in a manner that best achieves the objectives of policy. By contrast, exercising pure discretion gives the policymaker less influence over private-sector beliefs and, as a result, produces outcomes that are inferior to commitment.

Given the many advantages that commitment affords, it is surprising that the empirical literature has had little to say about which of the two policy concepts describes behavior that is closer to how central banks manage interest rates in the real world. Numerous papers analyzing US monetary policy, for example, assume that the Federal Reserve has access to some form of commitment technology, overlooking the possibility that discretion is more compatible with the data. Obscuring the issue further are past statements from leading figures in the policy-making community that offer different perspectives on central banking. In an article summarizing a 2007 speech by Philadelphia Fed President Charles Plosser, Michael Dotsey (2008, p. 8) asserts, "*The current Chairman, Ben Bernanke, is maintaining their* [Volcker and Greenspan's] *example of commitment to low and stable inflation. The benefits of following a committed plan are now so entrenched in policy-making circles that most central banks aggressively strive to maintain their credibility.*" Expressing an alternative view of Fed policy, Ben Bernanke and Frederic Mishkin (1997, pp. 105-113) remark that "*inflation targeting as it is actually practiced contains a considerable degree of what most economists would define as policy discretion* ... [and] *that a major reason for the success of the Volcker-Greenspan Fed is that it has employed a policymaking philosophy, or framework, which is de facto very similar to inflation targeting.*" Conflicting anecdotal accounts like these together with insufficient statistical evidence on the subject led McCallum (1999, p. 1489) to conclude "*that neither of these two modes of central bank behavior–rule-like* [commitment] *or discretionary–has as yet been firmly established as empirically relevant.*"

This paper attempts to bridge the gap in the research on commitment and discretion put forward by McCallum (1999). It starts from the presumption that the Federal Reserve sets interest rates in a deliberate fashion with specific goals in mind, and then asks whether it is possible to infer from the data which mode of optimization best explains macroeconomic outcomes in the US. Of course, discriminating commitment-like from discretionary activity over the sample can only be accomplished with an econometric procedure that gives voice to the explicit optimization problem of the policymaker. To that end, this paper borrows from a recent literature that estimates the parameters of an aggregate demand and supply model subject to the constraint that the policy component minimizes the central bank's loss function (e.g., Ozlale, 2003; Favero and Rovelli, 2003; Dennis, 2006). Conditioning estimation on the requirement that interest rates are chosen optimally, be it under commitment or discretion, enables one to obtain joint estimates of the structural parameters that characterize private behavior and the loss function weights that reveal the preferences of monetary policy.

To resolve which policy concept is the more "empirically relevant," I perform joint estimation separately under commitment and discretion and consider various measures of fit as a way of assessing congruence between the data and the models. The emphasis on relative *model* fit is appropriate because the two policies impose different cross-equation restrictions on the complete model in equilibrium. Utilizing all pertinent data within a coherent framework that accounts for the particular restrictions implied by commitment or discretion should help identify which policy is more likely to have produced the observed outcomes.

The exercise described above is carried out using a simple new-Keynesian model that governs the dynamics of output and inflation. The equations comprising this model form the constraints for the central bank's optimization problem. The stabilization goals of policy are represented by a quadratic loss function that penalizes deviations of inflation and output from target in addition to changes in the instrument setting. The last argument is often referred to as an interest rate smoothing objective. Structural parameters and loss function weights are estimated simultaneously, once under commitment and once under discretion, using a full information maximum-likelihood procedure with quarterly US data spanning the chairmanships of Volcker, Greenspan, and Bernanke. The series used for estimation include the output gap, inflation, and the nominal interest rate, variables that are likely to be informative about the conduct of monetary policy.

In evaluating the empirical performance of commitment and discretion, I appeal to formal measures of fit provided through the likelihood function as well as informal comparisons of the second moments implied by the estimated models. Regarding the latter criteria, I find that discretion does a better job of matching all the standard deviations calculated from the sample. It also dominates commitment in terms of the accuracy of the autocorrelations for inflation and the cross correlations between the output gap and inflation. For comparisons based on log likelihood, I employ the Bayesian information criterion and a related pseudo-posterior odds ratio which summarizes the probability of a model given the available data. I focus on these statistics rather than the raw log-likelihood values because the set of policies examined in this paper are non-nested. It turns out that the information criterion is considerably higher in the case of discretion, indicating greater fit with the data. The corresponding pseudo-odds measure reports a conditional probability of less than one percent for the commitment model compared to ninety-nine percent for the discretionary model. Taken together, these results point firmly to discretion as the preferred model of optimal policy.

Another issue examined in this paper concerns the extent to which commitment and discretion yield different estimates of key parameters, specifically the loss function weights. For each parameter I find that the estimates are similar across policies with one exception. Under commitment the weight on interest rate smoothing is significantly larger than the ones on inflation or the output gap. Under discretion, however, the weight on policy smoothing is estimated to be the smallest of the three, a more plausible result considering the traditional focus among central banks on the other two objectives (e.g., the "dual mandate" of the Federal Reserve). An inspection of the impulse response functions reveals that this divergence is mainly driven by the propensity for commitment to increase the volatility of the interest rate which, in turn, forces maximum likelihood to lift the smoothing penalty in an attempt to reconcile the model with the data.

The fact that empirical evidence strongly favors discretion raises the question of whether the US economy would have evolved differently had the Fed operated under commitment during the Volcker-Greenspan-Bernanke era. To shed light on this matter, I simulate the model with commitment using parameter values that correspond to the estimates obtained under discretion. The shocks used to generate the counterfactual series are the "true" structural shocks estimated from the discretionary model. Simulation results make clear that while the interest rate would have been more volatile under commitment, the paths of inflation and the output gap would have been similar to actual outcomes. As summarized by the loss function, the improvement in macroeconomic stabilization that would have occurred had policy been set under commitment rather than discretion is equivalent to a permanent shift in inflation of only 0.44 percentage points. The potential gains from commitment appear even smaller when the simulation takes into account the zero bound on nominal interest rates. Had a policy that respected the zero bound been enforced, interest rate volatility would have declined and the profiles for inflation and the output gap would have been even closer to their historical counterparts. The loss differential between commitment and discretion in this case is equivalent to a change in average inflation of 0.37 percentage points.

This paper is not the first to estimate a macroeconomic model subject to the auxiliary condition that monetary policy minimizes an explicit loss function. The earliest examples include Ozlale (2003) and Favero and Rovelli (2003), both of which estimate versions of the purely backward-looking model of Rudebusch and Svensson (1999). Employing different estimation strategies (Ozlale uses maximum likelihood whereas Favero and Rovelli use GMM), both papers find evidence of a structural break in the loss function weights after Volcker's appointment to chairman of the Federal Reserve. Dennis (2006) uses a similar model to estimate the Fed's implicit inflation target during the Volcker-Greenspan period and to examine whether the hypothesis of optimal policy can be formally rejected by the data. This study differs from the early literature in one important regard. The model used for estimation is forward-looking, implying a separation between optimal commitment and discretion that is altogether absent in the Rudebusch-Svensson model.

There is a recent literature that estimates central bank preferences in the context of forward-looking models. Dennis (2004) estimates a new-Keynesian model with discretionary policy and verifies that a structural break in the loss function occurred at the time of Volcker's appointment. Applying the same optimality restrictions, Söderström, Söderlind, and Vredin (2005) examine whether a loss function and a structural model can be jointly parameterized to match the broad moments in the US data. Castelnuovo (2006) demonstrates that adding forward-looking terms to a model with discretionary policy reduces the weight on interest rate smoothing needed to fit the data. Salemi (2006) examines the case of commitment to an optimized Taylor-type rule and concludes that the Fed placed greater emphasis on stabilizing inflation after 1980. Givens and Salemi (2008) employ the same model but focus instead on testing the efficiency of a proposed GMM strategy for joint estimation. Finally, Ilbas (2010a) uses Bayesian methods to estimate the euro-area model of Smets and Wouters (2003) conditional on the assumption that policy is set under full commitment.

While the papers in this literature deal with a variety of specific issues, they all share one aspect in common. Each paper makes an a priori assumption about the nature of monetary policy by considering only one of the two possible styles of optimization. In this study I take a step back and attempt to infer from the data which style is more empirically relevant. To my knowledge, this is the first paper that systematically compares the empirical effects of commitment and discretion by estimating the two policies side-by-side.

2 A Small Empirical Model of the US Economy

The model has three components. The first is an IS equation and a Phillips curve that govern the joint dynamics of output and inflation. Together, they form the constraints for the central bank's control problem. The second is a loss function that summarizes the stabilization goals of monetary policy. The third component is a procedure for determining the path of the nominal interest rate, namely, optimization under commitment or discretion.

2.1 The Policy Constraints

The constraints belong to a broad family of new-Keynesian models that have been applied extensively in the study of optimal monetary policy. There is a large literature showing that the aggregate behavioral equations comprising these models can be formally derived from an underlying dynamic general equilibrium framework (e.g., Kimball, 1995; Yun, 1996; Rotemberg and Woodford, 1997; McCallum and Nelson, 1999). The policy implications of a purely forward-looking class of new-Keynesian models are examined in detail by Clarida *et al.* (1999). The version used in this paper augments the basic forward-looking specification with various backward-looking elements designed to capture persistent aspects of the data (e.g., Fuhrer and Moore, 1995; Fuhrer, 1997; Estrella and Fuhrer, 2002).

It is important to acknowledge from the outset that while the policy constraints resemble the log-linear equations from an equilibrium model with particular structural features (which are discussed in more detail below), none of the exercises carried out in this paper directly concern estimation of the primitive factors representing the tastes and technologies of optimizing agents (although they can be readily applied in these settings). Instead, the emphasis is on estimation of a simple two-equation model describing local movements in output and inflation that is only partially consistent with general equilibrium. What this means is that the variables contained in the aggregate relationships are the same ones that appear endogenously in some optimization-based models, but the coefficients do not necessarily identify the deep, structural parameters that would be of primary interest in the latter.[1] Similar two-equation models with imperfect micro-foundations have recently been

[1]The coefficients of structural equations that evolve directly from general equilibrium are often complicated functions of the deep parameters describing the exact non-linear model. This paper does not take up the issue of identifying these primitive factors during estimation, which would require imposing numerous additional cross-parameter restrictions on the policy constraints.

estimated by Lindé (2005), Salemi (2006), and Cho and Moreno (2006). Jensen (2002), Walsh (2003), Ehrmann and Smets (2003), and Dennis and Söderström (2006) also use the same kind of models to evaluate the performance of alternative monetary policies.

Denote y_t the output gap, the log deviation of real output from potential, and let π_t be the inflation rate between dates $t-1$ and t. The output gap is determined by an IS equation

$$\begin{aligned} y_t &= \phi E_t y_{t+1} + (1-\phi)(\beta y_{t-1} + (1-\beta) y_{t-2}) - \sigma(i_t - E_t \pi_{t+1}) + u_{y,t}, \quad (1) \\ 0 &\leq \phi \leq 1,\ \beta \geq 0,\ \sigma > 0, \end{aligned}$$

where i_t is the (short-term) nominal interest rate, and E_t is an expectations operator conditional on date-t information. When $\phi = 1$, (1) corresponds to the log-linearized Euler equation describing the household's optimal consumption plan after imposing the resource constraint relating output to consumption. The parameter σ can be interpreted in this case as the intertemporal elasticity of substitution measuring the impact of changes in the real interest rate on current output. The exogenous variable $u_{y,t}$ is modeled as a mean-zero, serially uncorrelated demand shock (e.g., government spending) with constant variance σ_y^2.

The presence of lagged output gaps in the IS equation whenever $\phi < 1$ is largely motivated by empirical concerns. Estrella and Fuhrer (2002) and Fuhrer and Rudebusch (2004) assert that multiple lags are needed to explain the inertial responses of output observed in the data. Including lagged terms is also not inherently at odds with economic theory. Fuhrer (2000) shows that output persistence follows directly from first principles when the primitive model exhibits habit formation in consumption. An IS equation with the same lead-lag structure as (1) can be derived from a general equilibrium model that views habit formation as external to household optimization (e.g., Smets and Wouters, 2003; Ravn, Schmitt-Grohé, and Uribe, 2006) and dependent on exactly two lags of average consumption.[2]

[2] Refer to Appendix A for a derivation.

The inflation rate is governed by an expectations-augmented Phillips curve

$$\begin{aligned} \pi_t &= \alpha E_t \pi_{t+1} + (1-\alpha)\pi_{t-1} + \kappa y_t + u_{\pi,t}, \\ 0 &\leq \alpha \leq 1, \ \kappa > 0, \end{aligned} \tag{2}$$

which relates inflation to past and expected future inflation and the current output gap. For the case of $\alpha = 1$, (2) approximates the "new Phillips curve" estimated by Galí and Gertler (1999).[3] The new Phillips curve is consistent with a model of monopolistically competitive firms that stagger prices according to Calvo (1983). When opportunities to re-set prices occur, firms maximize a stream of profits subject to restrictions on the likelihood of subsequent price changes. The key cyclical factor affecting the pricing decision is real marginal cost, which can be shown to vary proportionately with the output gap (e.g., Woodford, 2003a, Ch. 3).[4] The slope coefficient κ carries information regarding the frequency of price revisions. Greater nominal rigidity, meaning less frequent revisions, implies a smaller value of κ. The stochastic variable $u_{\pi,t}$ is a mean-zero, serially uncorrelated supply shock with variance σ_π^2. It is often interpreted as a "cost-push" shock reflecting variations in marginal cost that do not shift the output gap (e.g., Clarida *et al.*, 1999). I allow for nonzero correlation between supply and demand shocks and denote their covariance $\sigma_{y\pi}$.

To account for the degree of inflation persistence found in the data, the Phillips curve includes a lagged term whenever $\alpha < 1$. Fuhrer and Moore (1995), Fuhrer (1997), and Roberts (1997) argue that without sufficient lagged dependence, the model produces "jump" dynamics that contradict empirical evidence suggesting inflation responds sluggishly to economic shocks. The presence of a backward-looking term can be motivated in theory by assum-

[3] The baseline Phillips curve presented in Galí and Gertler (1999) takes the form $\pi_t = \tilde{\beta} E_t \pi_{t+1} + \kappa y_t$. The parameter $\tilde{\beta}$ is the household's subjective discount factor and is estimated to be close to one.

[4] Galí and Gertler (1999) present evidence showing that the synthetic output gap variables used in most empirical studies (i.e., deterministic trends, indicators of capacity utilization, or the Congressional Budget Office estimate of potential GDP) are poor proxies for the theoretically consistent measure of real marginal cost based on the labor income share.

ing that prices charged by non-optimizing firms are automatically indexed to past inflation. Christiano, Eichenbaum, and Evans (2005) provide an example of full indexation, whereas the model presented in Smets and Wouters (2003) allows for partial indexation. Alternatively, lagged inflation can result from the existence of a group of firms that always forgo optimization in favor of using a simple rule-of-thumb for setting prices that depends on the recent history of competitors' prices (e.g., Galí and Gertler, 1999). Although it is not generally implied by partial indexation or rule-of-thumb behavior, restricting the coefficients on lagged and future inflation to sum to one ensures that (2) is conformable with the view that monetary policy has no long-run effect on output.[5]

2.2 The Loss Function

The central bank sets the path of the nominal interest rate to minimize the loss function

$$\mathbb{L}_t = E_t(1-\delta)\sum_{j=0}^{\infty}\delta^j\{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2\}, \tag{3}$$

where the discount factor $\delta \in (0,1)$ and $\lambda_y, \lambda_{\Delta i} \geq 0$. The first two terms penalize squared deviations of inflation and output from their respective targets. The inflation target is assumed to be constant over time and, without loss of generality, is normalized to zero. The target for output is the potential level.[6] The third term penalizes deviations of the interest rate from its previous level and is viewed as an interest rate smoothing incentive for the

[5]The partial indexation setup used by Smets and Wouters (2003) implies a Phillips curve of the form $\pi_t = \frac{\tilde{\beta}}{1+\gamma\tilde{\beta}}E_t\pi_{t+1} + \frac{\gamma}{1+\gamma\tilde{\beta}}\pi_{t-1} + \kappa(\theta,\gamma,\tilde{\beta})y_t$, where $\tilde{\beta}$ is the discount factor, γ is the degree of indexation to past inflation, and $1-\theta$ is the probability of price adjustment. The coefficients on lagged and future inflation sum to one when $\tilde{\beta}=1$ or $\gamma=1$ (full indexation). The Phillips curve derived from the rule-of-thumb model of Galí and Gertler (1999) takes the form $\pi_t = \frac{\tilde{\beta}\theta}{\theta+\omega[1-\theta(1-\tilde{\beta})]}E_t\pi_{t+1} + \frac{\omega}{\theta+\omega[1-\theta(1-\tilde{\beta})]}\pi_{t-1} + \kappa(\theta,\omega,\tilde{\beta})y_t$, where ω is the fraction of rule-of-thumb firms. Only when $\tilde{\beta}=1$ or $\omega=1$ do the inflation coefficients sum to one.

[6]The model is equivalent to one that accounts for a nonzero inflation target, but where the variables appearing in (1) and (2) are written as deviations from target values. The output target is still the potential level, and the nominal interest rate target is the sum of the inflation target and the steady-state real interest rate. See Dennis (2004) for an illustration of this equivalency.

policymaker. The weights λ_y and $\lambda_{\Delta i}$ measure the relative preference for stabilizing output and the interest rate smoothing argument.[7] The weight on inflation is normalized to one.

The loss function (3) is an appealing way to model central bank preferences for several reasons. First, parameterized versions of (3) are commonly used to assess the performance of different monetary policy rules (e.g., Rudebusch and Svensson, 1999; Levin and Williams, 2003). As a result, the estimates of λ_y and $\lambda_{\Delta i}$ have a familiar interpretation and are directly comparable to the broader literature on optimal monetary policy. Second, Svensson (1999) argues that the main objectives of a flexible inflation-targeting central bank can be described with a loss function that stabilizes inflation and a measure of real activity. This point is particularly relevant because many have argued that the Federal Reserve under Volcker and Greenspan employed a policy framework that closely resembles inflation targeting (e.g., Bernanke and Mishkin, 1997; Goodfriend, 2003). Third, Woodford (2002) shows that under certain conditions a loss function similar to (3), but without interest rate smoothing, is proportional to a quadratic approximation of the representative household's expected utility, so it provides a natural welfare criterion for ranking alternative policies.

Despite its theoretical appeal, one implication of using a strict, utility-based measure of loss is that the weights attached to the various objectives would depend on the underlying model parameters, notably the average frequency of price adjustments. By contrast, this paper regards the loss function weights as free parameters that are to be estimated jointly with the coefficients in the policy constraints. For this reason, and because it also contains an interest rate argument, the objective function given by (3) should not be interpreted as an approximation of the utility of a representative agent.

A few more comments on the limited theoretical foundations of (3) are in order. As

[7]The term "interest rate smoothing" is used throughout this paper to denote an explicit policy preference for reducing the variance of the nominal interest rate in first differences. There are other papers that interpret this phrase differently to mean a characteristic of the policy reaction function in which the coefficient on the lagged interest rate is substantially larger than zero (e.g., Levin, Wieland, and Williams, 1999).

illustrated by Woodford (2002), loss functions derived from first principles incorporate certain cross-equation restrictions with the model. Indeed, the technological assumptions that permit a quadratic approximation of expected utility to be written in terms of the variances of inflation and the output gap are the same ones that lead to a completely forward-looking version of the type of model described in (1) - (2). When lags are introduced into the IS equation through habit formation or into the Phillips curve through indexation or rule-of-thumb behavior, (3) fails to provide an accurate representation of consumer welfare even when $\lambda_{\Delta i} = 0$. Amato and Laubach (2004) show that habit formation generates a role for stabilizing actual output in addition to the output gap and inflation. In a model containing endogenous inflation persistence, Steinsson (2003) demonstrates that the loss function includes auxiliary terms reflecting the variances of the lagged output gap and the first difference of inflation. Because the structural equations estimated in this paper are not defined at the level of a general equilibrium model, I eschew formal micro-foundations of the policy objectives, adopting instead the more traditional loss function given by (3).

Finally, including an interest rate smoothing term, while difficult to justify in theory, is empirically compelling because it helps capture the degree of policy gradualism observed in the data. There are many explanations for why gradualism is desirable. Woodford (2003b) shows that in forward-looking models interest rate inertia is a defining feature of an optimal inflation-targeting rule. Brainard (1967) demonstrates that policy interventions should be cautious so as to avoid excess volatility resulting from misperceptions of a model with parameter uncertainty. Orphanides (2003) argues that similar caution is advisable when there is uncertainty regarding the accuracy of incoming data. Lowe and Ellis (1997) point out that such preferences may reflect a concern for financial market stability.

2.3 Optimal Monetary Policy

To compute optimal policies, stack the constraints in matrix form as

$$\begin{bmatrix} X_{t+1} \\ \Omega E_t x_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + Bi_t + \begin{bmatrix} \Gamma u_{t+1} \\ \mathbf{0}_{2\times 1} \end{bmatrix}, \tag{4}$$

where $X_t = [u_{y,t}\ u_{\pi,t}\ y_{t-1}\ y_{t-2}\ \pi_{t-1}\ i_{t-1}]'$ is a vector of predetermined variables, $x_t = [y_t\ \pi_t]'$ is a vector of forward-looking variables, and $u_{t+1} = [u_{y,t+1}\ u_{\pi,t+1}]'$ is a vector of innovations with covariance matrix Σ_{uu}. Structural parameters appear as elements of the matrices Ω, A, and B, and Γ is a selector matrix. The methods outlined by Söderlind (1999) are used to solve for the equilibrium dynamics under both commitment and discretion.[8]

The central bank selects the path of the nominal interest rate to minimize (3) subject to (4). Because the model is forward looking, policymakers face constraints that depend on expectations about the current and future course of monetary policy. In this kind of environment, the ability (or lack thereof) to credibly commit to a particular sequence of actions has important consequences on macroeconomic outcomes.

Under commitment the central bank announces at some specific date a complete, state-contingent plan for the interest rate that is to be strictly followed in all subsequent periods. When determining the path of optimal policy, it takes into account how the promise to execute such a contingency plan impacts private-sector expectations. In other words, the central bank internalizes the effect of its decisions on expected future variables in solving the optimization problem. Commitment thus presumes an ability to fulfill past promises and an understanding on the part of private agents of a willingness to do so regardless of what events transpire over time. This strategic interaction produces an optimal equilibrium in which the central bank makes efficient use of private-sector beliefs to achieve the goals

[8]Appendix B presents the matrices Ω, A, B, Γ, and Σ_{uu} for the model described in (1) - (2). It also provides details on the method used to solve for the equilibrium under commitment and discretion.

embodied by the loss function.[9]

It can be shown that the equilibrium law of motion under commitment follows

$$\begin{bmatrix} X_{t+1} \\ \psi_{t+1} \end{bmatrix} = M_c \begin{bmatrix} X_t \\ \psi_t \end{bmatrix} + \begin{bmatrix} \Gamma u_{t+1} \\ \mathbf{0}_{2\times 1} \end{bmatrix}, \tag{5}$$

$$\begin{bmatrix} x_t \\ i_t \end{bmatrix} = G_c \begin{bmatrix} X_t \\ \psi_t \end{bmatrix}, \tag{6}$$

where ψ_t is a vector of Lagrange multipliers associated with the lower block of (4) containing the forward-looking variables. Woodford (2003a, Ch. 7) explains that the multipliers capture the effect of expectations about the current policy setting that are reflected in the decisions private agents have made in all previous periods. It follows that actions taken by the central bank at any given time will depend on the full history of the economy dating back to the policy's inception. To see this more formally, note that one can solve for ψ_t in terms of the sequence $\{X_j\}_{j=0}^{t-1}$ by inverting the lag polynomial implied by the lower block of (5). Substituting the resulting expression into (6) gives a policy rule that describes the interest rate as a function of all current and past realizations of the predetermined variables.

Under discretion the central bank has the freedom to adjust policy in response to prevailing conditions. In contrast to commitment, that response does not have to be the one dictated by some contingency rule designed in an earlier period. Instead, a discretionary optimizer evaluates the current and prospective state of the economy and sets policy optimally on the basis of this assessment alone. It repeats this procedure each time a policy action is considered without ever making commitments regarding future choices. Because it cannot shape private-sector expectations in the absence of commitment, expected future variables

[9] Although the commitment policy is *ex ante* optimal, policymakers face an *ex post* incentive to abandon it (e.g., Kydland and Prescott, 1977). I show how to deal with the time-inconsistent nature of the commitment equilibrium as an empirical matter in the next section.

are taken as given in solving the optimization problem. The resulting equilibrium is only optimal in a constrained sense because the central bank, through sequential reoptimization, fails to harness expectations in a way that advances its stabilization goals.

The equilibrium law of motion under discretion is given by

$$
\begin{aligned}
X_{t+1} &= M_d X_t + \Gamma u_{t+1}, & (7)\\
\begin{bmatrix} x_t \\ i_t \end{bmatrix} &= G_d X_t. & (8)
\end{aligned}
$$

An important characteristic of the policy equation embedded in (8), and one that distinguishes it from commitment, is that it is purely forward looking rather than history dependent. A forward-looking policy is one in which outcomes are determined solely by the current and expected future outlook for the state of the economy. In equilibrium the interest rate depends only on today's predetermined variables since conditional expectations are computed within the model as linear projections onto the current state.

3 Estimation Strategy

The recursive equilibrium under commitment or discretion takes the form of an empirical state-space model that can be estimated with maximum likelihood using the Kalman filtering algorithms described in Hamilton (1994, Ch. 13). The state equation is

$$\xi_{t+1} = F\xi_t + \tilde{\Gamma} u_{t+1}, \tag{9}$$

where $\xi_t \in \{[X_t' \ \psi_t']', \ X_t\}$, $F \in \{M_c, \ M_d\}$, and $\tilde{\Gamma} \in \{[\Gamma' \quad \mathbf{0}_{2\times 2}]', \ \Gamma\}$. The observation equation is

$$\begin{bmatrix} y_t^o \\ \pi_t^o \\ i_t^o \end{bmatrix} = H\xi_t + \begin{bmatrix} 0 \\ 0 \\ v_{i,t} \end{bmatrix}, \tag{10}$$

where $\{y_t^o, \ \pi_t^o, \ i_t^o\}_{t=1}^T$ denotes the observed series for the output gap, inflation, and the nominal interest rate, and $H \in \{G_c, \ G_d\}$. The stochastic variable $v_{i,t}$ is a mean-zero, serially uncorrelated shock to the observed interest rate. Its variance σ_i^2 is interpreted as a measure of the discrepancy between the optimal policy implied by the model and the actual interest rate in the sample. With data on three variables, adding $v_{i,t}$ also circumvents the stochastic singularity problem emphasized by Ingram, Kocherlakota, and Savin (1994).

The structural parameters are estimated with quarterly US data over the period 1982:Q1 - 2008:Q4.[10] The output gap is the log deviation of real GDP from potential GDP as constructed by the Congressional Budget Office. Inflation is the first difference of the log of the GDP implicit price deflator expressed at an annual rate. The nominal interest rate is the annual percentage yield on 3-month Treasury bills.

Prior to estimation the inflation and interest rate series are de-meaned so that their sample averages match the mean values of the corresponding data generated by the model. Because no constant intercept terms appear in either the IS equation (1) or the Phillips curve (2), the model describes a mean-zero process for all variables in the system. De-meaning also implies that the target levels for inflation and the interest rate implicit in the loss function (3) are equal to the averages taken from the data. Such an assumption is useful for two reasons. First, the main objective is to compare model fit and estimates of the loss function weights under commitment and discretion rather than obtain estimates of the Fed's latent target

[10]Following Dennis (2006), the sample begins in 1982:Q1 in order to exclude the period when the Federal Reserve's operating procedure focused on targeting the quantity of non-borrowed reserves.

values.[11] Second, de-meaning ensures that the target levels for inflation and the interest rate are the same for each mode of policy optimization considered during estimation.

There is some evidence suggesting that inflation and the nominal interest rate are potentially non-stationary in US data.[12] To see whether the sample employed in this paper contains unit roots, I perform a series of augmented Dickey-Fuller tests using lag lengths zero through four. One can reject the null hypothesis of a unit root in inflation at the one-percent significance level with lag lengths zero through two and at the ten-percent level with lags three through four. Concerning the interest rate, one can reject the hypothesis of non-stationarity using a lag length of one at the ten-percent level, but the test fails to reject the unit-root hypothesis for all other lag selections.[13] These last results should be interpreted with caution because unit-root tests are known to have low power to reject the null in small samples, especially when the data are highly serially correlated. In what follows I regard the interest rate as a mean-reverting process for the purpose of estimation.

Application of the Kalman filter typically begins with a date-0 estimate of the initial state vector, call it $\hat{\xi}_{1|0}$, equal to its long-run mean. The mean value consistent with (9) is zero. This raises a potential concern for estimation under commitment because ξ_t contains the Lagrange multipliers associated with the forward-looking variables in (4). Starting the recursion with $\psi_t = 0$ implies that policy decisions are unconstrained in the initial period, so any actions taken in that period do not have to confirm private-sector expectations that were formed at dates preceding the beginning of the sample. As a result, the central bank will find it optimal to implement the discretionary policy just once in the first period while

[11] This approach has recently been used by Ozlale (2003), Söderström *et al.* (2005), and Castelnuovo (2006). Papers that estimate the inflation target directly include Favero and Rovelli (2003), Dennis (2004), Dennis (2006), Ireland (2007), and Fève, Matheron, and Sahuc (2010).

[12] Barsky (1987), Ball and Cechetti (1990), and Culver and Papell (1997) report that the hypothesis of a unit root in inflation cannot be rejected in US data. Perron (1988), Stock and Watson (1988), and Mishkin (1992) find similar evidence for the nominal interest rate.

[13] Starting with four lags, the Dickey-Fuller regressions are pared down by using standard t-tests until the last lag is significantly different from zero. This procedure points to one lag as the correct number to use in conducting a test on the interest rate and three lags as the correct number for inflation.

promising to behave in a committed fashion thereafter. This aspect of the commitment equilibrium, known as time inconsistency, is fundamental to the optimal control of forward-looking systems (e.g., Kydland and Prescott, 1977), but it is problematic for estimation due to the arbitrary significance it places on the first observation of the sample. By initializing the Kalman filter with $\psi_t = 0$, the model's interpretation of past events would be one in which the Federal Reserve ignored any commitments made prior to 1982:Q1 while committing to a new plan that was optimal from the standpoint of that specific date onward. This would undoubtedly be viewed as a shortcoming of the analysis since there is no compelling historical evidence to indicate that such a regime change took place on that particular date.

In estimating the commitment model, it is better to assume that the Fed has announced its contingency rule at some unspecified point predating the sample. It follows that the economy's initial evolution (in the first period of the sample) will be consistent with policy actions taken at all dates after the starting period. This is equivalent to adopting an equilibrium concept that relates closely to what Woodford (2003a, Ch. 7) calls the optimal policy from a "timeless perspective." A timeless perspective policy requires that central bank actions always validate previously-formed expectations even in the initial period. In practice it is found by substituting out the Lagrange multipliers from the first-order conditions of the control problem, yielding a time-invariant criterion involving only loss function variables that must be satisfied in every period (e.g., Giannoni and Woodford, 2005; Dennis, 2010).

By contrast, I allow the multipliers to enter the state-space model, but I initialize the Kalman filter with nonzero values so that policy does not arbitrarily deviate from the commitment program at the beginning of the sample. The implementation of the commitment equilibrium during estimation is perhaps more similar to the optimal policy examined in Khan, King, and Wolman (2003). To make the commitment problem fully recursive and the resulting policy time consistent, the authors treat all forward-looking constraints as strictly binding in the initial period. This requires augmenting the standard Lagrangian with a sufficient number of lagged multipliers, one for each new constraint imposed on the optimization problem. The inclusion of auxiliary multipliers makes the first-order conditions time invariant, implying the same policy behavior at all dates over the planning horizon.

Although some elements of ξ_t are unobservable to the econometrician, an informed decision about their starting values can be made using the Kalman filter. The strategy employed here is to estimate the model in a first stage by setting $\hat{\xi}^{(1)}_{1|0} = 0$, at which point the Kalman filter is used to generate a sequence of date-t updated projections of the state $\{\hat{\xi}^{(1)}_{t|t}\}_{t=1}^{T}$. The model is then re-estimated taking as the initial state the mean value of forecasts computed in the previous step, that is, $\hat{\xi}^{(2)}_{1|0} = (1/T)\sum_{t=1}^{T}\hat{\xi}^{(1)}_{t|t}$. This process is repeated until the initial state equals the average forecast produced by the Kalman filter, or $\hat{\xi}^{(i+1)}_{1|0} = \hat{\xi}^{(i)}_{1|0}$. For estimation under commitment and discretion, convergence occurred in less than 10 iterations.[14]

4 Empirical Results

4.1 Maximum-Likelihood Estimates

Table 1 displays the point estimates and standard errors of the model's 11 structural parameters.[15] The left panel presents estimates for the case of commitment, and the right panel presents estimates for discretion. The standard errors are computed by taking the square roots of the diagonal elements of the information matrix, obtained by inverting the matrix of second derivatives of the maximized log-likelihood function.[16]

There are numerous similarities in the estimates across policies. Looking first at the

[14] An alternative strategy adopted by Ilbas (2010a) is to partition the sample to include an initialization period that precedes estimation. She finds that a presample period of 20 quarters is sufficient to eliminate any effects on parameter estimates of setting the multipliers equal to zero in the initial period. Estimating the model in this paper using a presample initialization period produced results similar to the ones reported using the iterative approach described above.

[15] The central bank's discount factor δ is set equal to 0.99 prior to estimation. Stress tests showed that the remaining estimates are not overly sensitive to variations in this parameter.

[16] Appendix C plots conditional forecasts of the commitment multipliers generated by the Kalman filter and discusses the significance of the observed variation over the sample.

covariances, estimates of σ_y and σ_π indicate that supply shocks are more than twice as volatile as demand shocks. Estimates of σ_i suggest that the disparity between observed interest rates and the optimal ones implied by the model are largely invariant to the two modes of central bank behavior.

Turning next to the IS and Phillips curves, estimates of ϕ and α indicate that forward and backward-looking terms are important for output and inflation dynamics. The estimate of ϕ is close to one-third under commitment and about 0.37 under discretion. Both values are in the neighborhood of those reported by Fuhrer and Rudebusch (2004) and Lindé (2005). The estimates of α are 0.49 and 0.62, echoing the findings of Roberts (2005) and Kiley (2007). Estimates of κ, the output gap elasticity of inflation, are small but within the range typical of the literature. Values of κ near zero are evidence of long duration of price stickiness or large strategic complementarities accompanied by modest nominal rigidities.[17]

Regarding the loss function, estimates of λ_y point to a small preference for output gap stability relative to inflation. There is some recent empirical work suggesting that the Fed demonstrates little concern for output stability as an independent goal of policy. In a model featuring an optimized Taylor-type rule estimated over the period 1980:Q1 - 2001:Q4, Salemi (2006) finds that the relative weight on output stabilization is only 0.0012 and not significantly different from zero.[18] In a related model but with policy set under discretion, Dennis (2004) reports estimates equal to zero for both the Volcker-Greenspan period (1983:Q1 - 2002:Q2) and for the subperiod covering only Greenspan's chairmanship (1987:Q3 - 2002:Q2). Dennis (2006), in the context of a backward-looking model spanning 1982:Q1 - 2000:Q2, also concludes that the estimate of λ_y is not significantly different from zero.

To see if the weight on output gap volatility is significant in the present model, I conduct likelihood ratio tests of the null hypotheses that $\lambda_y = 0$ under commitment and discretion.[19]

[17]See Woodford (2003a, Ch. 3) for details.

[18]Givens and Salemi (2008) report similar findings using the same model estimated with GMM rather than maximum likelihood.

[19]An alternative test of parameter significance is based on the Wald statistic, formed by squaring the

The relevant test statistic is formed by doubling the difference between the unrestricted and restricted log-likelihood values. It is asymptotically distributed under the null as a chi-square random variable with one degree of freedom.

The first two columns of Table 2 display estimates and log-likelihood values associated with each policy when λ_y is held fixed at zero. A comparison between Table 1 and Table 2 reveals that most of the parameters are not greatly affected by the absence of an output gap objective. Despite broad similarities in the estimates, the likelihood ratio statistics are 4.18 (p-value is 0.041) in the commitment case and 8.14 (p-value is 0.004) in the discretion case. The hypothesis that $\lambda_y = 0$ is, therefore, rejected by the data under both policy specifications.[20] These findings are consistent with evidence reported in Favero and Rovelli (2003), Ozlale (2003), and Ilbas (2010b) showing that the preference for output stability is quantitatively small but statistically significant. Favero and Rovelli (2003), for example, estimate λ_y to be 0.00125 with a standard error of 0.0002 for the period 1980:Q3 - 1998:Q3.

Returning to Table 1, estimates of $\lambda_{\Delta i}$ make clear that the revealed preference for interest rate smoothing depends on the nature of optimal policy. Under commitment the estimate of $\lambda_{\Delta i} = 2.56$, indicating that central bank attitudes are chiefly concerned with smoothing interest rates followed by inflation and then output stability. This result is similar to Dennis (2004), Dennis (2006), and Söderström *et al.* (2005) who contend that optimal and historical policy can be reconciled provided interest rate smoothing appears as the dominant objective in the Fed's loss function. These authors obtain statistically significant estimates of $\lambda_{\Delta i}$ ranging from 1.109 to 4.517 over dates that fall within the Volcker-Greenspan era. A very different outcome emerges under discretion, where the estimate of $\lambda_{\Delta i} = 0.06$ places the

ratio of the point estimate to its standard error. Its dependence on the latter means that the Wald statistic can be distorted by approximation error that results from numerical computation of the Hessian and its corresponding inverse matrix. When estimates of the restricted model are available, the likelihood ratio test is preferred because it can be conducted without reference to standard errors.

[20]A Wald test of the hypothesis that $\lambda_y = 0$ cannot be rejected in the discretionary model. Appendix D explains why the Wald test contradicts the likelihood ratio test in this case.

interest rate objective just below output in the central bank's ordering of policy goals. Studies that find evidence of a weak preference for interest rate smoothing include Favero and Rovelli (2003), Castelnuovo and Surico (2004), and Ilbas (2010b). Using a variety of different models and estimation techniques, these authors report estimates of $\lambda_{\Delta i}$ spanning 0.0085 to 0.16.

To evaluate the significance of interest rate smoothing, I re-estimate the parameters subject to separate restrictions on $\lambda_{\Delta i}$ under commitment and discretion. In the commitment case I fix $\lambda_{\Delta i} = 1$ prior to estimation. A likelihood ratio test of this restriction amounts to a test of the hypothesis that inflation and interest rate smoothing receive the same weight in the Fed's loss function. In the discretion case I set $\lambda_{\Delta i} = 0.01$. While the intent was to test the hypothesis that $\lambda_{\Delta i} = 0$, preliminary attempts to estimate the model failed because the forecast error covariance matrix of the data became rank-deficient. Setting $\lambda_{\Delta i} = 0$ ensures that policymakers fully insulate the output gap and inflation from demand shocks because there is no direct penalty for adjusting the interest rate. With supply shocks as the only remaining source of exogenous variation, the model implies an exact deterministic relationship between the output gap and inflation that cannot be reconciled with the data.[21]

The final two columns of Table 2 display estimates and log-likelihood values for the restrictions on $\lambda_{\Delta i}$ described above. Setting $\lambda_{\Delta i} = 1$ under commitment has a significant effect on estimates of the other parameters. The estimates of σ_y, ϕ, and σ, for example, point to much larger demand shocks in the IS equation, a diminished role for expected future output, and a stronger real interest rate channel. Given the size of the standard errors, however, considerable uncertainty remains about the true values of these parameters. Imposing $\lambda_{\Delta i} = 1$ also lowers maximized log likelihood from -387.77 to -396.56, producing a likelihood ratio statistic equal to 17.57 (p-value < 0.001). The hypothesis that inflation

[21]Due to numerical inaccuracies in computing log likelihood, values of $\lambda_{\Delta i}$ close to zero can produce a covariance matrix for the data that is nearly singular. Based on numerous estimation trials, I found that $\lambda_{\Delta i} = 0.01$ was small enough to form inferences about the statistical contribution of interest rate smoothing but not so small as to risk encountering econometric problems resulting from stochastic singularity.

and interest rate smoothing receive the same weight in the loss function is, therefore, rejected with a high degree of confidence in the commitment model. By contrast, fixing $\lambda_{\Delta i} = 0.01$ under discretion has little impact on parameter estimates and only changes log likelihood from -379.61 to -379.68. The implied likelihood ratio statistic is 0.16 (p-value is 0.694), so the hypothesis that $\lambda_{\Delta i}$ is a small number close to zero cannot be rejected in the model with discretionary policy. I return to the issue of why inferences about the role of interest rate smoothing are so different across the two policies in a later section.

4.2 Model Comparison

This section compares the empirical performance of commitment and discretion using two criteria. First, a group of second moments are calculated from the historical data and compared to ones generated by the estimated models to see how well each policy captures key features of the US business cycle. Classical estimation does not focus exclusively on matching this limited set of moments, so a second comparison is made by appealing to a broad measure of fit provided through the likelihood function. Specifically, the Bayesian information criterion is computed for both policies as well as a corresponding posterior odds measure that reveals the probability of a model given the data.

Table 3 presents the standard deviations of inflation, the output gap, and the interest rate as implied by the data and the models. The model with discretionary policy does a better job of accounting for the standard deviations of all three variables. The volatility of inflation, in particular, is only slightly larger than the realized volatility in the data. In comparison, the commitment model significantly overstates inflation and output gap volatility.

Figure 1 plots vector autocorrelation functions for the same variables. As in Fuhrer and Moore (1995), the autocorrelations for the data are computed from a fourth-order vector autoregression. Most of the correlations produced by the estimated models match closely their counterparts from the data. For example, both commitment and discretion deliver substantial output gap persistence as measured by correlations between current and lagged output gaps. The half-life of this autocorrelation is about 5 quarters. A similar result applies to the persistence of the nominal interest rate. The models also account for the positive and declining lead-lag relationship between inflation and the interest rate.

There are two areas where discretionary policy generates a more visible improvement in fit. The first is the degree of inflation persistence reflected in the autocorrelation function for inflation. The half-life of this function is 2 quarters under discretion but about 5 quarters under commitment. Discretionary policy also improves the accuracy of the cross correlations between the output gap and inflation. In the absence of commitment, the model correctly predicts the sign and magnitude of this relationship at leads and lags of up to one year.

Another way to assess model fit is with the Bayesian information criterion (BIC). The BIC is a consistent model-selection criterion that penalizes likelihood by an amount that increases with the number of estimated parameters. An advantage of the BIC is that it facilitates the comparison among a group of non-nested models. Estimation under commitment and discretion results in a pair of non-nested models because neither policy can be obtained by placing certain parametric restrictions on the other.[22] The BIC for model j is

$$BIC(j) = \log \mathcal{L}(j) - \frac{N(j)}{2} \ln(T),$$

where $\log \mathcal{L}(j)$ is the maximized value of log likelihood, $N(j)$ is the number of estimated parameters in model j, and T is the sample size.

The BIC can also be used to form a pseudo-posterior odds ratio that gives the data-determined probability of a model. Kiley (2007) explains that in large samples the BIC approximates the marginal likelihood of a model in which the data, summarized by the likelihood function, predominates the Bayesian prior distribution of the parameters. A pseudo-

[22]Minimizing the loss function under commitment and discretion imposes separate coefficient restrictions on the policy equation of a model that, in equilibrium, takes the form of a trivariate VAR.

odds measure is then formed by replacing marginal likelihood with the BIC in the ratio

$$\rho(j) = \frac{\exp(BIC(j))}{\sum_{h=1}^{z} \exp(BIC(h))},$$

where $\rho(j)$ is the conditional probability of model j among the family of z different models under consideration. Although it is consistent with a Bayesian approach to model selection, the pseudo-odds ratio is determined solely by the quality of the model's characterization of the data (with an adjustment for degrees of freedom) and not by any prior information concerning the parameter or model space. This follows directly from the implicit use of equal prior model probabilities in the construction of $\rho(j)$ and from the BIC being invariant to priors over the parameters within each model.[23]

Table 4 reports log likelihood, the BIC, and the pseudo-posterior odds ratio for the models under commitment and discretion.[24] The BIC is -413.47 for commitment but -405.31 for discretion. As a result, the pseudo-odds statistic points to a very small conditional probability of 0.0003 in the commitment model compared to 0.9997 in the discretionary model.

The evidence presented in favor of discretion leaves open the question of whether the data support the hypothesis that Federal Reserve behavior was optimal during the Volcker-Greenspan-Bernanke period. Comparisons of fit between commitment and discretion alone are insufficient to answer this question because both models assume that policy is set optimally, with the only difference being in how policymakers manage expectations. To arrange a valid test of the optimal-policy hypothesis, the comparison group must include an encompassing model that does not constrain central bank actions to be the outcome of a loss minimization problem. To that end, I re-estimate (1) and (2) jointly with an unrestricted

[23] Recent studies that utilize the BIC for model comparison include Brock, Durlauf, and West (2003), Kiley (2007), and Keen (2009).

[24] As a criterion for selecting between commitment and discretion, the BIC is equivalent to maximum likelihood because the number of parameters are equal in the two models.

equation for the interest rate that attaches separate response coefficients to every variable in the state vector, namely, demand and supply shocks, one lag each of inflation and the interest rate, and two lags of the output gap. The model with discretionary policy is a special case of this three-equation system that results from conditioning estimation on the assumption that the response coefficients minimize an expected loss function.

Relaxing the coefficient restrictions implied by optimal discretion results in the interest rate equation (with standard errors in parentheses)

$$i_t = \underset{(0.0605)}{1.8651} u_{y,t} + \underset{(0.0653)}{0.1389} u_{\pi,t} + \underset{(0.0795)}{0.5533} y_{t-1} - \underset{(0.0818)}{0.4836} y_{t-2} + \underset{(0.0320)}{0.2982} \pi_{t-1} + \underset{(0.0183)}{0.8821} i_{t-1}. \tag{11}$$

The interest rate equation produced by the discretionary model is

$$i_t = \underset{(0.1919)}{2.3612} u_{y,t} + \underset{(0.0448)}{0.5099} u_{\pi,t} + \underset{(0.0658)}{0.7088} y_{t-1} - \underset{(0.0581)}{0.6618} y_{t-2} + \underset{(0.0230)}{0.1957} \pi_{t-1} + \underset{(0.0100)}{0.9017} i_{t-1}, \tag{12}$$

where standard errors are found using the delta method. Because the models are nested, one can test the hypothesis that policy was set under discretion by computing the likelihood ratio statistic and comparing it to the appropriate critical value from the chi-square distribution.[25] Maximized log likelihood in the model containing the unrestricted policy rule (11) is -371.36 and is reported in the third row of Table 4. This model has 15 free parameters, the 9 structural parameters describing (1), (2), and the covariance matrix of the shocks plus 6 distinct policy-rule coefficients. The discretionary model, which returns a log likelihood of -379.61, has the same 9 structural parameters but only 2 free parameters in the loss function. It follows that the optimal rule given by (12) places a total of 4 restrictions on the response coefficients estimated in (11). Under the null that (12) is true, the likelihood ratio statistic is asymptotically distributed as a chi-square random variable with 4 degrees of freedom. The

[25]The null model is specified in terms of discretion because previous results indicated that it dominates commitment according to several measures of empirical fit.

actual likelihood ratio statistic is 16.50 (p-value < 0.001), so the data reject the hypothesis that historical outcomes were the result of discretionary optimization according to this test.

Using pseudo-Bayesian analysis for model comparison leads to a different conclusion about which policy fits the data best. The BIC for the three-equation model without optimal policy is -406.40, which is slightly smaller than the corresponding value for discretion but larger than the one for commitment. When all three are considered, the implied pseudo-posterior odds ratios displayed in Table 4 indicate a conditional probability of about 75 percent for the discretionary model compared to only 25 percent for the unrestricted model. The probability of the commitment model is near zero. The pseudo-odds criterion, therefore, points to discretion as the preferred model given the available data.

The conflicting evidence on model fit provided by the classical likelihood ratio test and the pseudo-odds criterion appears to be reconciled by the correction for degrees of freedom accounted for in the BIC. It is clear from Table 4 that loosening the coefficient restrictions implied by optimal discretion improves the model's raw characterization of the data as measured by gains in log likelihood. From a pseudo-Bayesian perspective, however, that improvement is evidently not large enough given the sample size to offset the loss of degrees of freedom that results from introducing several new parameters into the model.

4.3 The Role of Interest Rate Smoothing

Inferences about the role of interest rate smoothing as an independent policy goal vary greatly depending on the mode of optimization. Estimates under commitment suggest that it is the most important objective in the Fed's loss function, but discretion implies that it is the least. This section provides intuition for why maximum likelihood produces contradictory findings.

Figure 2 graphs the impulse responses to demand and supply shocks for two different versions of the model. The first version is the estimated model under discretion. The second version takes the same parameters (including loss function weights) but assumes that policy is set under commitment. The ensuing differences in model dynamics are, therefore, driven entirely by the procedure for determining interest rates. As expected, commitment leads to more stable dynamics of the output gap and inflation. Consider first the effects of a positive demand shock. Both policies generate plausible "hump-shaped" movements in the output gap and inflation, but the amplitude and persistence of the responses are smaller under commitment. Inflation jumps by the same amount under both policies following a positive supply shock, but the steady-state reversion is more gradual under discretion. It is easy to see how the central bank achieves greater stability by examining the responses of the interest rate. By committing in advance to keep interest rates elevated for an extended length of time, policymakers reduce expectations of future output gaps and inflation and, consequently, dampen the adjustment of those variables in the near term. The only exception is in the response of the output gap to a supply shock, in which case the period of high interest rates prescribed by commitment results in a sustained drop in output below potential. The scale of this departure is much smaller than the one induced by a demand shock, so on balance the total volatility of the output gap is still lower under commitment.[26]

The problem with the commitment outcome from an empirical perspective is that it implies a level of interest rate volatility that is severely at odds with the data. Any data-fitting exercise will naturally seek parameter values that drive down this volatility. Figure 3 illustrates some of the key tradeoffs faced by maximum likelihood when locating the estimate of $\lambda_{\Delta i}$. The figures plot standard deviations of inflation, the output gap, and the interest rate for a range of values of $\lambda_{\Delta i}$, holding the other parameters fixed at their point estimates. The left panel corresponds to the discretionary model, and the right panel corresponds to the commitment model. The vertical lines indicate the estimates of $\lambda_{\Delta i}$ in each model, and the crosses identify the sample moments that are reported in the first column of Table 3.

In the case of discretion, raising $\lambda_{\Delta i}$ lowers the standard deviation of the interest rate but

[26]Refer to Table 5 in the next section for confirmation of this result.

has little effect on output and inflation. Most of the reduction occurs at small values of $\lambda_{\Delta i}$. Interest rate volatility is extremely high for values near zero but below historical levels for $\lambda_{\Delta i}$ above 0.09. A much bigger weight is needed to reconcile the model with the data under commitment. For $\lambda_{\Delta i}$ near the discretionary estimate, the interest rate is twice as volatile as the actual series. Increasing $\lambda_{\Delta i}$ causes the standard deviation of the interest rate to fall but those of the output gap and inflation to rise. As a result, maximum likelihood compromises between these moments by selecting a large value of $\lambda_{\Delta i}$ at which the volatility of all three variables exceed their sample counterparts by nontrivial amounts.

The weight on interest rate smoothing is not the only reason why inflation and output gap volatility are too high under commitment. As Figure 3 makes clear, the standard deviations of both variables are larger than those from the data even when $\lambda_{\Delta i} = 0$. The explanation is straightforward. During the course of estimation, maximum likelihood concentrates on an area of the parameter space that strikes a balance between all of the moments embodied by the likelihood function. It turns out that the relevant area is one that implies less forward-looking (more backward-looking) behavior in the IS and Phillips curves. All else equal, smaller estimates of ϕ and α increase the persistence and volatility of output and inflation by putting more weight on the lagged terms in (1) and (2). Söderström *et al.* (2005) demonstrate the same property in a related new-Keynesian model with optimal policy.

Figure 4 illustrates this point by graphing the response functions for the models estimated with commitment and discretion. In contrast to Figure 2, changes in dynamics are now driven by the policy specification as well as variation in the structural parameters. As a consequence of the large estimate of $\lambda_{\Delta i}$, the interest rate responses to demand and supply shocks under commitment are much closer to the adjustment paths observed under discretion. Combined with smaller estimates of ϕ and α, greater concern for policy smoothing leads to an increase in the volatility of the output gap and inflation. The peak effect of a demand shock on inflation is twice as large for commitment and occurs three quarters later than discretion.

Supply shocks cause inflation to rise in both models, but the adjustment back to steady state is more gradual under commitment. The output gap responses exhibit a similar pattern of increased volatility when moving from discretion to commitment.

5 Counterfactual Analysis

Macroeconomic outcomes over the sample period are more consistent with the notion that the Federal Reserve set policy under discretion rather than commitment. Not only does the discretionary model fit the data better, inferences about the relative importance of interest rate smoothing accord with traditional views on the primary goals of US monetary policy. These findings motivate the following retrospective question. Assuming that historical policy actions were the result of discretion, how much better off would outcomes have been had the Fed precommitted to an optimal rule? This section performs counterfactual simulations of the model estimated with discretionary policy to ascertain how the economy might have evolved under commitment. Plots of the simulated data reveal the extent to which these monetary factors alone could have altered the historical profiles of output and inflation.

Figure 5 displays the actual series for inflation, the output gap, and the interest rate and the paths these variables would have taken had the Fed implemented the commitment policy from 1982:Q1 to 2008:Q4. To produce counterfactual data, the fixed interval Kalman smoother described by de Jong (1989) is used to estimate the history of shocks implied by the discretionary model.[27] The shocks are reinserted back into the model, holding the parameters constant (including loss function weights) but with the policy shifted to commitment.[28]

Simulations reveal that the path of the interest rate would have been different under com-

[27]The smoothed estimates of the shocks reflect information contained in the full set of data.

[28]In a related exercise Dennis (2005) estimates shocks from a baseline model that describes policy with a forward-looking Taylor rule. He goes on to simulate two distinct sets of counterfactual data by feeding the shocks into a variant of the baseline model that replaces the Taylor rule with a calibrated loss function. One simulation assumes that policy minimizes loss under commitment, and the other assumes discretion.

mitment. Rates would have fallen rapidly during the 1980s, bottoming out at 1.16 percent in 1986:Q4 before reaching a high of 7.46 percent in 1990:Q4 and staying close to historical levels for the rest of the decade. After peaking at 10.62 percent in 2000:Q4, the interest rate would have tumbled for the duration of the sample. Overall, the volatility of the simulated path is much larger than what actually transpired. The maximum difference between the counterfactual and observed series is 6.59 percentage points, occurring in 2002:Q4.

Despite very different policy behavior at times, the paths of inflation and the output gap would have been remarkably similar to historical outcomes. Inflation would have been slightly smaller before 1995:Q1 under commitment and slightly larger thereafter. The biggest gap between the two series is only 0.53 percentage points in 1990:Q2. The output gap also would have been somewhat lower in the late 1980s and early 1990s but a bit higher in the second half of the 1990s and again after 2002:Q1.

Although useful for historical comparisons, counterfactual simulations do not easily translate into a single measure that quantifies the cumulative losses associated with one policy relative to another. For this purpose I follow Jensen (2002) and Dennis and Söderström (2006) and compute the "inflation equivalent," interpreted as the permanent increase in inflation from target that in terms of central bank loss is equivalent to moving from commitment to discretion. The inflation equivalent can be calculated from (3) as $\pi^{eq} = \sqrt{\mathbb{L}_d - \mathbb{L}_c}$, where $\mathbb{L}_d$ and $\mathbb{L}_c$ are the losses under discretion and commitment, respectively.[29] A similar quantity measured in terms of lost output, an output gap equivalent, is given by $y^{eq} = \pi^{eq}/\sqrt{\lambda_y}$.

The top panel of Table 5 reports loss under both policies in addition to the inflation and output gap equivalents from discretion. It also presents the variances of inflation, the output gap, and the interest rate in first differences and levels. Moving from discretion to commitment leads to small reductions in the variances of inflation and the output gap and a

[29] A permanent inflation rate of π^{eq} yields a loss equal to $(1-\delta)\sum_{j=0}^{\infty}\delta^j \pi^{eq^2} = \pi^{eq^2}$. The inflation equivalent, therefore, satisfies $\mathbb{L}_c + \pi^{eq^2} = \mathbb{L}_d$.

comparable increase in the variance of interest rate changes. The variance of the interest rate in levels is considerably larger under commitment. Loss is about 10 percent higher under discretion, a gap that is equivalent to a permanent increase in inflation of 0.44 percentage points or an output gap of 1.40 percentage points.

The modest improvement in macroeconomic stabilization overstates the true gains that commitment could have achieved once one considers that such a policy would not have been operational. The reason is that the level of interest rate volatility induced by commitment virtually assures frequent violations of the zero lower bound. Indeed, Figure 5 reveals that the simulated path of the interest rate becomes negative in 2004:Q4 and continues to decline through 2008. As a result, the contingency rule derived under commitment does not describe a policy that the central bank could have actually implemented.

To reassess the gains from commitment while respecting the presence of the zero bound, I follow Woodford (2003a, Ch. 6) in approximating the effects of this constraint by amending the loss function to include a penalty on the variance of the interest rate. The idea is that by placing a large enough weight on the auxiliary term, the probability that the nonnegativity constraint ever binds can be made small.[30] In this exercise I select a value for the weight on interest rate variability just sufficient to ensure that the counterfactual series is nonnegative at every date over the sample period.[31] The results are illustrated in Figure 6.

The interest rate path would still have been lower than the actual series through most of the 1980s, but the pace of monetary easing would have been slower with a zero bound constraint. The simulated data bottoms out at 3.80 percent in 1988:Q1 and remains near historical levels during the 1990s. Rates would have declined more gradually after 2000:Q4, but they also would have stayed low after the Fed began to tighten in 2004:Q1. The largest

[30] Eggertsson and Woodford (2003) and Adam and Billi (2007) study optimal monetary policy subject to the restriction that the interest rate is nonnegative in every state. Replacing this exact constraint with an alternative one on interest rate variability conserves the linear-quadratic setup of the model.

[31] A value of $\lambda_i = 0.0024$ in a loss function of the form $E_0(1-\delta)\sum_{t=0}^{\infty}\delta^t\{\pi_t^2 + \lambda_y y_t^2 + \lambda_{\Delta i}(i_t - i_{t-1})^2 + \lambda_i i_t^2\}$ guarantees a strictly positive counterfactual sequence of nominal interest rates.

gap between the counterfactual and observed series is 3.76 percentage points, occurring in 2002:Q4. Not surprisingly, inflation and the output gap would have been even closer to their historical counterparts had policy accounted for the zero bound. The biggest differences in the inflation and output series are 0.48 and 0.73 percentage points (1990:Q3 and 2004:Q4).

The bottom panel of Table 5 shows how these outcomes translate into loss. Moving from discretion to commitment while respecting the zero bound generates even smaller reductions in the variances of inflation and the output gap. It also lowers the variance of the interest rate in first differences but increases it in levels. Loss under discretion is about 6 percent higher than commitment, yielding an inflation equivalent of only 0.37 percentage points or an output gap equivalent of 1.17 percentage points.

6 Sensitivity Analysis

This section examines whether the estimates and results concerning model fit are robust to certain changes in the data and the model. The changes include re-estimation with data on the Consumer Price Index and a measure of "core" inflation, adding an interest rate variance argument to the loss function, and specifying policy in terms of an optimized simple rule.

6.1 Using CPI Data for Estimation

The first two columns of Table 6 report estimates under commitment and discretion for the case in which the Consumer Price Index (CPI) replaces the GDP implicit price deflator as the index for constructing inflation. Because it is published monthly, I convert the CPI into a quarterly series by computing segmented three-month averages beginning in January-March of 1982 and ending in October-December of 2008. Inflation is the first difference of the log of the resulting series expressed at an annual rate.

Incorporating CPI data leads to significant changes in some parameters. Compared to the values in Table 1, estimates of the standard deviation of the supply shock σ_π more than double under commitment and discretion when the dataset includes CPI inflation. The same is true of the slope coefficient κ in the Phillips curve, the estimates of which increase by an order of magnitude under both policies. Regarding the loss function, estimates of λ_y and $\lambda_{\Delta i}$ using CPI data reveal a diminished concern for output stability in both models but a larger concern for interest rate smoothing. Notwithstanding these changes in the policy weights, interest rate smoothing remains the leading stabilization objective in the commitment model while inflation dominates the loss function in the discretionary model. Finally, likelihood-based measures of fit indicate that discretion still provides a better summary of the time-series properties of the data. Log likelihood and the BIC are substantially larger for discretion and, consequently, the pseudo-posterior odds ratio attaches very little probability to commitment.

6.2 Using PCE Data for Estimation

The concept of "core" inflation has recently become a major topic of discussion within the Federal Reserve, and Congressional testimony of current and past FOMC members bears witness to this development (e.g., Wynne, 2008). Although there is no formal agreement on the ideal way to measure core inflation, the most common approach is to simply omit from an economy-wide index of prices certain items that experience volatile price movements such as food and energy products (e.g., Bernanke, Laubach, Mishkin, and Posen, 1999). Indeed, the headline measure of core inflation often referenced in published Federal Reserve transcripts is the annual percentage change in the deflator for Personal Consumption Expenditures (PCE) less food and energy prices.[32] Below I re-estimate the model using inflation data assembled from the core PCE deflator. The results are in the third and fourth columns of Table 6.

Most of the parameter estimates are not significantly altered by the use of PCE data. The

[32]The belief is that core inflation is a better guide to monetary policy than overall inflation because it is more informative about the underlying trend in prices.

obvious exceptions are the weights on output gap stability λ_y and interest rate smoothing $\lambda_{\Delta i}$, both of which increase considerably under commitment and discretion when the dataset includes core inflation. Despite these common shifts in the policy objectives, the amount of uncertainty surrounding λ_y and $\lambda_{\Delta i}$ differs sharply between the two policies. Under commitment the standard errors are small relative to the point estimates. Separate Wald tests of the hypothesis that each weight is zero can be rejected at normal significance levels. In the discretion case the standard errors are large by comparison, implying that the same tests of parameter significance cannot be rejected with any reasonable degree of confidence.

Regarding model fit, the selection criteria now firmly point to commitment as the preferred model when estimated with PCE inflation. This finding should be interpreted with great caution as it could simply be an indication that both models fail to adequately characterize the data on core inflation. In fact, to take seriously the idea that historical outcomes were the result of commitment in the context of the present model, one would have to accept the notion that policymakers viewed interest rate smoothing as eleven times more important than inflation stability. To put this number into perspective, note that $\lambda_{\Delta i} = 11.61$ implies that the loss function penalizes a 1-percentage-point increase in inflation the same as a 0.29-percentage-point change in the interest rate. Such a strong aversion to interest rate volatility seems implausible considering the belief among academics and central bankers alike that monetary policy has traditionally placed more emphasis on inflation and the output gap in accordance with the so-called "dual mandate" of the Federal Reserve (e.g., Mishkin, 2007).

6.3 A Loss Function with Interest Rate Stabilization

There is some dispute in the literature about the appropriate way to incorporate a preference for interest rate stability in the loss function. The assumption made in this paper and in many others is that the central bank penalizes variation in the first difference of the interest rate, a behavior commonly referred to as interest rate smoothing. By contrast, others define stability to mean low variation in the policy instrument around a fixed target level (e.g., Giannoni and Woodford, 2005; Salemi, 2006; Onatski and Williams, 2010).[33] In this section I allow the data to ascertain which view of interest rate stability is more empirically relevant by estimating an augmented loss function that encompasses both objectives. Specifically,

$$\mathbb{L}_t = E_t(1-\delta)\sum_{j=0}^{\infty}\delta^j\{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2 + \lambda_i i_{t+j}^2\}, \tag{13}$$

where $\lambda_i \geq 0$. The final term in (13) penalizes squared departures of the interest rate from a constant target, which is normalized to zero to preserve coherence between the model and the data. The results for commitment and discretion are in the last two columns of Table 6.

Adding an interest rate variance argument to the commitment model produces results that are identical to those in Table 1. During the course of estimation, λ_i converges to the lower bound of the admissible parameter space where the loss function reduces to the original specification given by (3). Augmenting the loss function also has little impact on estimates of the discretionary model. Most of the parameters are close to their counterparts in Table 1, and the estimate of λ_i is near zero. Restricting $\lambda_i = 0$ under discretion only lowers maximized log likelihood from -379.12 to -379.61. The implied likelihood ratio statistic equals 0.98 (p-value is 0.323), so the hypothesis that $\lambda_i = 0$ cannot be rejected by the data.

6.4 An Optimized Simple Rule

This section estimates a version of the model that characterizes monetary policy with a simple Taylor-type rule rather than commitment or discretion. The policy rule is simple in that it prescribes how the interest rate responds directly to changes in a particular subset of the state variables, making it easy for the central bank to communicate its actions to

[33] A quadratic loss function in which the interest rate appears in levels rather than differences can be motivated in theory by the presence of transactions frictions that give rise to a demand for money or, as discussed in the last section, a desire to reduce the likelihood of violating the zero lower bound (e.g., Woodford, 2003a, Ch. 6).

the public and for the public, in turn, to verify compliance with the rule. Following the advice of McCallum (1999), I assume that the only states featured in the rule are the lagged endogenous variables. Exogenous shocks and contemporaneous endogenous variables are excluded on the grounds that neither would be observable to actual policymakers when setting the current interest rate. The class of policy rules considered for estimation is

$$i_t = \theta_i i_{t-1} + (1 - \theta_i)(\theta_\pi \pi_{t-1} + \theta_{y1} y_{t-1} + \theta_{y2} y_{t-2}), \tag{14}$$

where θ_π, θ_{y1}, θ_{y2}, and θ_i are the coefficients that determine the responses to lagged observations of inflation, the output gap, and the interest rate.

In keeping with the theme of optimal policy, the central bank chooses $\Theta = \{\theta_\pi, \theta_{y1}, \theta_{y2}, \theta_i\}$ once and for all to minimize the loss function (3) subject to the aggregate behavioral equations (1) and (2). The role of policy thereafter is to implement the optimized rule in every period, recognizing that its prescriptions will be contingent on prevailing economic circumstances. When solving for Θ, the central bank takes into account how its commitment to carry out (14) at all future dates impacts private expectations.[34] The procedure for computing an optimal simple rule is straightforward. For a given set of policy-rule coefficients, find the rational expectations solution to the system made up of (1), (2), and (14) and use it to derive the asymptotic value of loss. This step can be accomplished with any standard linear solution method (e.g., Klein, 2000). Next, activate a hill-climbing algorithm that searches over the elements of Θ that reduce loss and stops when no smaller number can be found.[35]

Before discussing the results, I address an econometric issue concerning estimation of the standard errors. As noted in Salemi (2006), maximum-likelihood estimation of a model where policy takes the form of an optimal simple rule produces a log-likelihood function

[34] Clarida *et al.* (1999) and others refer to policies formulated in this manner as "commitment" to an optimal simple rule.

[35] The domain of Θ is restricted during estimation to ensure that the policy rule (14) is consistent with a determinate rational expectations equilibrium.

that is not differentiable around the maximum. Consequently, methods that depend on approximating the Hessian or the outer product of the score are incapable of generating meaningful standard errors. To provide a measure of statistical significance in the absence of reliable standard errors, Salemi reports p-values from separate likelihood ratio tests on every parameter.

An inspection of the likelihood surface for the present model reveals similar irregularities that make it impossible to obtain standard errors in the usual way. As an alternative to the Salemi approach, I adopt a Bayesian strategy to quantify the uncertainty regarding parameter estimates. The procedure involves coupling the likelihood function with a prior distribution over the parameters using Bayes' theorem to form a joint posterior distribution from which the sampling variability of the estimates can be inferred. Following Onatski and Williams (2010), I formulate independent uniform prior densities for the 11 structural parameters. The bounded ranges over which the priors are defined permit a large area of the parameter space to be explored when constructing the posterior. To generate a sequence of draws from the posterior, I employ the random walk Metropolis-Hastings algorithm described in An and Schorfheide (2007). The draws are used to approximate the posterior mean and 95 percent confidence interval for each parameter.[36] The findings are reported in Table 7.

Also listed in Table 7 are the posterior mode estimates, found by maximizing the sum of the log-likelihood function and the log-prior distribution of the parameters.[37] In this case the posterior mode estimates are identical to estimates one would obtain by simply maximizing log likelihood over the support of the prior. This follows directly from the use of uniform priors in evaluating the posterior density. Because it views all points within the support as equally probable and those outside as having zero probability, a uniform prior does not inform the data in any significant way. It merely truncates the range of values deemed

[36]Details concerning estimation of the posterior density can be found in Appendix E.

[37]The posterior mode is the element of the parameter space that maximizes the posterior density, which according to Bayes' rule is proportional to the product of the likelihood function and the prior distribution.

permissible for maximum-likelihood estimation.[38]

The parameters most affected by the switch to an optimal simple rule are the ones in the IS equation. Inferences about the size of demand shocks σ_y roughly double when moving from commitment or discretion to a simple rule. This shift appears to be statistically significant as estimates of σ_y under the original policies lie outside the posterior 95 percent probability interval. Inferences about the interest rate elasticity σ and the coefficient on expected future output ϕ also change considerably under a simple rule. The maximum-likelihood (posterior mode) estimate of the former increases to 0.05 while that of the latter drops to 0.14. These results are similar to ones reported in Salemi (2006). Using an optimized simple rule identical to (14), Salemi estimates σ and ϕ to be 0.042 and 0.162, respectively.

Regarding loss function parameters, the estimate of $\lambda_y = 0.10$ is basically unchanged from the discretionary model. Draws from the posterior indicate a high degree of precision in this estimate, with the 95 percent confidence interval spanning 0.06 to 0.14. Turning to interest rate smoothing, the estimate of $\lambda_{\Delta i} = 0.76$ falls between the values obtained under commitment and discretion. The posterior distribution of $\lambda_{\Delta i}$ is also concentrated tightly around the mean, suggesting that the data are highly informative about this parameter.

The central bank's control problem defines Θ as an implicit function of the structural parameters and loss function weights. The optimized rule for the estimates in Table 7 is

$$i_t = 0.8924 i_{t-1} + (1 - 0.8924)(3.0527\pi_{t-1} + 8.9221 y_{t-1} - 7.9016 y_{t-2}). \tag{15}$$

Reconciling an optimal simple rule with the data evidently requires large countercyclical responses to inflation and the output gap in the long run but only gradual adjustment of the interest rate to this desired level in the short run.

[38]None of the posterior mode estimates presented in Table 7 fall near the edges of the prior density, indicating that classical and Bayesian estimation are equivalent in the present context. Incorporating priors, therefore, serves only to facilitate estimation of the posterior means and 95 percent probability intervals.

I conclude this section with a brief discussion about model fit. Among the three types of optimal policy estimated in this paper, discretion results in the highest log-likelihood value followed by commitment and then the simple rule. Pseudo-odds ratios constructed from the BIC still point to discretion as the dominant model, assigning near zero probability to the other two models. The relative performance of the optimal simple rule also highlights the importance for model fit of allowing policy to respond contemporaneously to economic shocks. A key difference between commitment and discretion on the one hand and the policy rule (14) on the other is that the latter specification restricts the policymaker's information set to include only the lagged state variables. When monetary policy conditions on an expanded state that accommodates current demand and supply shocks, log likelihood improves by nearly 8 points in the case of commitment and by 16 points in the case of discretion.

7 Concluding Remarks

This paper reports estimates from a new-Keynesian model of the US economy in which monetary policy minimizes the central bank's loss function. The forward-looking nature of the model draws a distinction between two modes of optimization, termed commitment and discretion. The model is estimated separately under each policy using quarterly data over the Volcker-Greenspan-Bernanke era. The goal is to judge which form of central bank optimization fits the data best and to determine the extent to which the two procedures generate different estimates of the structural parameters, particularly the loss function weights.

Maximum-likelihood estimates point to broad similarities in the parameters across policies with one major exception. The weight on interest rate smoothing is large under commitment but small under discretion. The result can be traced to the fact that commitment increases the volatility of the interest rate, and maximum likelihood tries to offset this effect by lifting the weight on policy smoothing. Measures of fit based on the likelihood function indicate that discretionary policy provides a superior description of the joint time-series properties of the data. Additional evidence in favor of discretion can be drawn from comparisons of the standard deviations and autocorrelation functions implied by the two models.

The empirical analysis carried out in this paper uses a strictly binary framework in which the policy choices are either full commitment or discretion. In future work it might be more realistic to think about Fed behavior as lying somewhere between these two logical extremes. Schaumburg and Tambalotti (2007) develop a modeling device, which they call "quasi-commitment," that makes it possible to analyze a continuum of policies between commitment and discretion that differ in degree of credibility. In short, policymakers are understood by the public to renege on the optimal commitment plan every period with some constant probability. Outcomes converge to full commitment as this probability approaches zero and to discretion as it approaches unity. Using the quasi-commitment framework, it should be possible to estimate simultaneously the weights in the central bank's loss function and the exogenous probability that identifies it's measure of credibility.

Appendix A. An IS Curve with Two Lags of Output

Consider an economy populated by a large number of identical households that make intertemporal consumption/saving decisions and supply labor to the production sector in competitive markets. The preferences of the representative household are given by

$$E_0 \sum_{t=0}^{\infty} \tilde{\beta}^t \left\{ U\left(C_t - H_t; u_t\right) - \nu\left(L_t\right) \right\}, \tag{A.1}$$

where U is a monotonic and concave utility function defined over consumption C_t relative to habit consumption H_t. The consumption good is a Dixit-Stiglitz aggregate of intermediate products sold by monopolistically competitive firms. The variable u_t is a mean-zero, serially uncorrelated taste shock that exogenously shifts the marginal utility of income. The function ν is the disutility of supplying labor L_t and is strictly increasing and convex. The parameter $\tilde{\beta}$ is the subjective discount factor, and E_0 is a conditional (date-0) expectations operator.

Households value consumption relative to an external habit H_t that evolves according to

$$H_t = b_1 \overline{C}_{t-1} + b_2 \overline{C}_{t-2}, \tag{A.2}$$

where $\overline{C}_{t-j}$ is average consumption at $t-j$ and $|b_1 + b_2| < 1$. The restriction on (b_1, b_2) ensures that utility is increasing in steady-state consumption. Under external habit formation increases in average consumption boost the marginal utility of consumption in subsequent periods. Holding the marginal utility of income fixed, the optimal response is to raise individual consumption in those periods, producing a "catching up with the Joneses" effect.

The household's flow budget constraint takes the form

$$C_t + \frac{B_t}{P_t} \leq W_t L_t + \frac{I_{t-1} B_{t-1}}{P_t} + Div_t, \tag{A.3}$$

where P_t is the price of the consumption good, B_{t-1} denotes the quantity of riskless, one-period bonds carried into period t, and I_{t-1} is the corresponding gross nominal interest rate. $W_t L_t$ represents wage income, and Div_t is a stream of real profits from ownership of firms.

The representative household chooses $\{C_t, L_t, B_t\}_{t=0}^{\infty}$ by maximizing (A.1) subject to (A.2) and (A.3), taking as given the processes $\{\overline{C}_t, P_t, W_t, I_t, Div_t; u_t\}_{t=0}^{\infty}$ and the initial values $\overline{C}_{-1}$, $\overline{C}_{-2}$, B_{-1}, and I_{-1}. The first-order conditions are

$$U_c\left(C_t - b_1\overline{C}_{t-1} - b_2\overline{C}_{t-2}; u_t\right) = \Lambda_t, \tag{A.4}$$

$$\Lambda_t = \tilde{\beta} E_t\left[\Lambda_{t+1} I_t \frac{P_t}{P_{t+1}}\right], \tag{A.5}$$

$$W_t = \frac{\nu_L\left(L_t\right)}{\Lambda_t}, \tag{A.6}$$

where Λ_t is the Lagrange multiplier associated with (A.3).

To obtain an IS equation that links output Y_t to past and expected future output and the real interest rate, combine the log-linear approximations of (A.4) and (A.5) with the equilibrium requirement $Y_t = \overline{C}_t = C_t$. After some rearranging, the approximation becomes

$$\hat{Y}_t = \frac{1}{1+b_1} E_t\hat{Y}_{t+1} + \frac{b_1 - b_2}{1+b_1}\hat{Y}_{t-1} + \frac{b_2}{1+b_1}\hat{Y}_{t-2} - \frac{1-b_1-b_2}{\tilde{\sigma}(1+b_1)}\left(\hat{I}_t - E_t\hat{\Pi}_{t+1} - u_t\right), \tag{A.7}$$

where $\Pi_{t+1} \equiv P_{t+1}/P_t$ and any variable $\hat{Z}_t \equiv \log(Z_t/Z)$. The parameter $\tilde{\sigma} \equiv -(C-H)U_{cc}/U_c$ governs the curvature of the utility function with respect to $C_t - H_t$ in the steady state.

There are important similarities between (A.7) and the IS equation (1) in section 2.1. First, both equations have the same distributed lag structure. Current output depends on next period's expected output and output in the previous two periods. Second, the coefficients on the lead and lag terms sum to one, making output in the long run invariant to permanent shifts in the real interest rate. Third, both equations permit a negative coefficient on the second output lag. This occurs in (1) anytime $\beta > 1$ and in (A.7) when $b_2 < 0$.

Appendix B. Model Solution

The model described in (1) - (2) can be expressed in compact form as

$$\begin{bmatrix} X_{t+1} \\ \Omega E_t x_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ x_t \end{bmatrix} + B i_t + \begin{bmatrix} \Gamma u_{t+1} \\ \mathbf{0}_{2\times 1} \end{bmatrix}, \tag{B.1}$$

where $X_t = [u_{y,t}\ u_{\pi,t}\ y_{t-1}\ y_{t-2}\ \pi_{t-1}\ i_{t-1}]'$, $x_t = [y_t\ \pi_t]'$, and $u_{t+1} = [u_{y,t+1}\ u_{\pi,t+1}]'$ with covariance Σ_{uu}. The matrices Ω, A, B, Γ, and Σ_{uu} are given by

$$\Omega = \begin{bmatrix} \phi & \sigma \\ 0 & \alpha \end{bmatrix}, \quad A = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -(1-\phi)\beta & -(1-\phi)(1-\beta) & 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 & -(1-\alpha) & 0 & -\kappa & 1 \end{bmatrix},$$

$$B = [\,0\ \ 0\ \ 0\ \ 0\ \ 1\ \ \sigma\ \ 0\,]', \quad \Gamma = \begin{bmatrix} \mathbf{I}_{2\times 2} \\ \mathbf{0}_{4\times 2} \end{bmatrix}, \quad \Sigma_{uu} = \begin{bmatrix} \sigma_y^2 & \sigma_{y\pi} \\ \sigma_{y\pi} & \sigma_\pi^2 \end{bmatrix}.$$

Similarly, denote $S_t = [\pi_t\ y_t\ \Delta i_t]'$ the vector of target variables appearing in the loss function.[39] S_t is related to the state vector and the policy instrument by

$$S_t = D \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix}, \text{ where } D = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}.$$

[39] Δ is the first difference operator.

Reformulating (3) in terms of S_t, the loss function at date 0 can be written as

$$\mathbb{L}_0 = E_0(1-\delta)\sum_{t=0}^{\infty}\delta^t S_t' W S_t, \tag{B.2}$$

where $W = [1 \;\; \lambda_y \;\; \lambda_{\Delta i}] \times \mathbf{I}_{3\times 3}$ is a symmetric, positive-semidefinite matrix with diagonal elements corresponding to the policy weights.

B.1 The Commitment Equilibrium

Consider the problem of minimizing (B.2) under commitment subject to (B.1) for $t \geq 0$ and $X_0 = \overline{X}_0$, where $\overline{X}_0$ is known. Following Söderlind (1999), I construct the Lagrangian

$$\mathcal{L}_0 = E_0 \sum_{t=0}^{\infty}(1-\delta)\delta^t \left\{ S_t' W S_t + 2[\xi_{t+1}' \;\; \psi_{t+1}'] \left(\overline{\Omega} \begin{bmatrix} X_{t+1} \\ x_{t+1} \\ i_{t+1} \end{bmatrix} - \overline{A} \begin{bmatrix} X_t \\ x_t \\ i_t \end{bmatrix} - \begin{bmatrix} \Gamma u_{t+1} \\ \mathbf{0}_{2\times 1} \end{bmatrix} \right) \right\}$$
$$+ \frac{1-\delta}{\delta}\xi_0'(X_0 - \overline{X}_0),$$

where ξ_{t+1} and ψ_{t+1} are vectors of Lagrange multipliers associated with the upper and lower blocks of (B.1), respectively.[40] Note that the constraint in $\mathcal{L}_0$ has been restated so that the policy instrument i_t now appears as the final element in the state vector. This transformation requires defining auxiliary matrices $\overline{\Omega}$ and $\overline{A}$ as

$$\overline{\Omega} \equiv \begin{bmatrix} \mathbf{I}_{6\times 6} & \mathbf{0}_{6\times 2} & \mathbf{0}_{6\times 1} \\ \mathbf{0}_{2\times 6} & \Omega & \mathbf{0}_{2\times 1} \end{bmatrix}, \quad \overline{A} \equiv [\; A \;\; B \;].$$

[40] In forming the Lagrangian I have made use of the law of iterated expectations.

The first-order conditions with respect to X_t, x_t, and i_t are

$$[\ X_t' \ \ x_t' \ \ i_t'\ ]\, D'WD + [\ \xi_t' \ \ \psi_t'\ ]\, \frac{1}{\delta}\overline{\Omega} - [\ E_t\xi_{t+1}' \ \ \psi_{t+1}'\ ]\, \overline{A} = 0 \tag{B.3}$$

for $t \geq 0$, where $X_0 = \overline{X}_0$ and $\psi_0 = \mathbf{0}_{2\times 1}$. The last equality follows from the fact that there is no constraint associated with the lower block of (B.1) in the initial period.[41]

The first-order conditions (B.3) can be combined with the policy constraints (as they appear in the Lagrangian) to form a system of 17 difference equations for $t \geq 0$

$$\begin{bmatrix} \overline{\Omega} & \mathbf{0}_{8\times 8} \\ \mathbf{0}_{9\times 9} & \overline{A}' \end{bmatrix} \begin{bmatrix} X_{t+1} \\ E_t x_{t+1} \\ E_t i_{t+1} \\ E_t \xi_{t+1} \\ \psi_{t+1} \end{bmatrix} = \begin{bmatrix} \overline{A} & \mathbf{0}_{8\times 8} \\ D'WD & \frac{1}{\delta}\overline{\Omega}' \end{bmatrix} \begin{bmatrix} X_t \\ x_t \\ i_t \\ \xi_t \\ \psi_t \end{bmatrix} + \begin{bmatrix} \Gamma u_{t+1} \\ \mathbf{0}_{2\times 1} \\ \mathbf{0}_{9\times 1} \end{bmatrix}, \tag{B.4}$$

where X_t and ψ_t are the predetermined variables (i.e., variables with exogenous one-step-ahead forecast errors), and x_t, i_t, and ξ_t are the non-predetermined variables.

I follow Klein (2000) in using the generalized Schur form to separate (B.4) into stable and unstable blocks of equations. Given $X_0 = \overline{X}_0$ and $\psi_0 = \mathbf{0}_{2\times 1}$, a unique bounded solution exists if the number of stable eigenvalues equals the number of predetermined variables. Assuming the determinacy condition is satisfied, the solution takes the form

$$\begin{bmatrix} X_{t+1} \\ \psi_{t+1} \end{bmatrix} = M_c \begin{bmatrix} X_t \\ \psi_t \end{bmatrix} + \begin{bmatrix} \Gamma u_{t+1} \\ \mathbf{0}_{2\times 1} \end{bmatrix}, \tag{B.5}$$

$$\begin{bmatrix} x_t \\ i_t \end{bmatrix} = G_c \begin{bmatrix} X_t \\ \psi_t \end{bmatrix}, \tag{B.6}$$

[41] Because it is measurable with respect to the information set at date t, $E_t\psi_{t+1} = \psi_{t+1}$ in (B.3).

where the matrices M_c and G_c depend on Ω, A, B, W, and δ.[42]

B.2 The Discretion Equilibrium

Consider the problem of minimizing the date-t intertemporal loss function $\mathbb{L}_t$ under discretion subject to (B.1), X_t given, and

$$i_{t+1} = F_{t+1}X_{t+1}, \tag{B.7}$$

$$x_{t+1} = G_{t+1}X_{t+1}, \tag{B.8}$$

where F_{t+1} and G_{t+1} are the (unknown) decision rules that solve the minimization problem at date $t+1$. Following Söderlind (1999), I construct the Bellman equation associated with the optimal value of the loss function in period t. To transform the control problem into a standard optimal linear regulator program, it is necessary to rewrite the forward-looking variables x_t in terms of the predetermined states X_t and the control i_t.

First, taking conditional expectations of (B.1) gives

$$\begin{bmatrix} \mathbf{I}_{6\times 6} & \mathbf{0}_{6\times 2} \\ \mathbf{0}_{2\times 6} & \Omega \end{bmatrix} \begin{bmatrix} E_t X_{t+1} \\ E_t x_{t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} X_t \\ x_t \end{bmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} i_t, \tag{B.9}$$

where A and B have been partitioned into submatrices conformable with the dimensions of X_t and x_t. Next, using (B.8) and the upper block of (B.9) yields

$$E_t x_{t+1} = G_{t+1} E_t X_{t+1} = G_{t+1}(A_{11}X_t + A_{12}x_t + B_1 i_t). \tag{B.10}$$

Multiplying (B.10) by Ω and setting the resulting expression equal to the lower block of

[42]There is also a solution for ξ_t, but it is not needed to characterize the dynamics of X_t, ψ_t, x_t, or i_t.

(B.9) leads to a solution for x_t of the form

$$x_t = \tilde{A}_t X_t + \tilde{B}_t i_t, \tag{B.11}$$

where

$$\begin{aligned}\tilde{A}_t &\equiv (A_{22} - \Omega G_{t+1} A_{12})^{-1}(\Omega G_{t+1} A_{11} - A_{21}),\\ \tilde{B}_t &\equiv (A_{22} - \Omega G_{t+1} A_{12})^{-1}(\Omega G_{t+1} B_1 - B_2).\end{aligned}$$

Substituting (B.11) into the upper block of (B.9) and reinserting the innovations then gives

$$X_{t+1} = A_t^* X_t + B_t^* i_t + \Gamma u_{t+1}, \tag{B.12}$$

where

$$\begin{aligned}A_t^* &\equiv A_{11} + A_{12}\tilde{A}_t,\\ B_t^* &\equiv B_1 + A_{12}\tilde{B}_t.\end{aligned}$$

Finally, using (B.11) to express S_t in terms of X_t and i_t allows one to write period loss as

$$S_t' W S_t = \begin{bmatrix} X_t \\ i_t \end{bmatrix}' \begin{bmatrix} Q_t & N_t \\ N_t' & R_t \end{bmatrix} \begin{bmatrix} X_t \\ i_t \end{bmatrix}, \tag{B.13}$$

where

$$\begin{aligned}Q_t &\equiv \tilde{W}_{XX} + \tilde{A}_t' \tilde{W}_{Xx}' + \tilde{W}_{Xx}\tilde{A}_t + \tilde{A}_t'\tilde{W}_{xx}\tilde{A}_t,\\ N_t &\equiv \tilde{W}_{Xx}\tilde{B}_t + \tilde{A}_t'\tilde{W}_{xx}\tilde{B}_t + \tilde{W}_{Xi} + \tilde{A}_t'\tilde{W}_{xi},\\ R_t &\equiv \tilde{B}_t'\tilde{W}_{xx}\tilde{B}_t + \tilde{W}_{xi}'\tilde{B}_t + \tilde{B}_t'\tilde{W}_{xi} + \tilde{W}_{ii}.\end{aligned}$$

The substitutions resulting in (B.13) require defining an auxiliary matrix $\tilde{W}$ as

$$\tilde{W} \equiv D'WD = \begin{bmatrix} \tilde{W}_{XX} & \tilde{W}_{Xx} & \tilde{W}_{Xi} \\ \tilde{W}'_{Xx} & \tilde{W}_{xx} & \tilde{W}_{xi} \\ \tilde{W}'_{Xi} & \tilde{W}'_{xi} & \tilde{W}_{ii} \end{bmatrix},$$

which can be partitioned into submatrices conformable with X_t, x_t, and i_t.

Because the model is linear-quadratic, the optimal value of the loss function at date $t+1$ admits the quadratic form, $X'_{t+1}V_{t+1}X_{t+1} + \upsilon_{t+1}$, where V_{t+1} is a positive-semidefinite matrix and υ_{t+1} is a scalar independent of X_{t+1}. It follows that the value of the loss function in period t satisfies the Bellman equation

$$(1-\delta)\left[X'_tV_tX_t + \upsilon_t\right] = (1-\delta)\min_{i_t}\left\{S'_tWS_t + \delta E_t\left[X'_{t+1}V_{t+1}X_{t+1} + \upsilon_{t+1}\right]\right\} \tag{B.14}$$

subject to (B.12), (B.13), and X_t known.

The first-order condition with respect to i_t is

$$N'_tX_t + R_ti_t + \delta B^{*\prime}_tV_{t+1}(A^*_tX_t + B^*_ti_t) = 0. \tag{B.15}$$

Solving (B.15) for i_t gives

$$i_t = F_tX_t, \tag{B.16}$$

where

$$F_t \equiv -(R_t + \delta B^{*\prime}_tV_{t+1}B^*_t)^{-1}(N'_t + \delta B^{*\prime}_tV_{t+1}A^*_t). \tag{B.17}$$

Inserting (B.16) into (B.11) yields a solution for x_t of the form

$$x_t = G_tX_t,$$

where

$$G_t \equiv \tilde{A}_t + \tilde{B}_t F_t. \tag{B.18}$$

Finally, using (B.16) in (B.14) and matching the common terms gives

$$V_t \equiv Q_t + N_t F_t + F_t' N_t' + F_t' R_t F_t + \delta(A_t^* + B_t^* F_t)' V_{t+1}(A_t^* + B_t^* F_t). \tag{B.19}$$

The system of equations describing $\tilde{A}_t$, $\tilde{B}_t$, A_t^*, B_t^*, Q_t, N_t, R_t, F_t, G_t, and V_t constitutes a mapping from $(F_{t+1}, G_{t+1}, V_{t+1})$ to (F_t, G_t, V_t). The solution to the control problem is a fixed point (F, G, V) of this mapping and is obtained as the limit of (F_t, G_t, V_t) when $t \to -\infty$. Assuming that the fixed point problem is stable, the discretion equilibrium takes the form

$$X_{t+1} = M_d X_t + \Gamma u_{t+1}, \tag{B.20}$$

$$\begin{bmatrix} x_t \\ i_t \end{bmatrix} = G_d X_t, \tag{B.21}$$

where $M_d = A^* + B^* F$ and $G_d = [\ G' \quad F'\ ]'$.

Appendix C. The Commitment Multipliers

Two variables that affect equilibrium dynamics under commitment are the Lagrange multipliers attaching to the IS equation (1) and the Phillips curve (2). Grouped in the vector ψ_t, the commitment multipliers measure the change in loss that results from a temporary relaxation of the forward-looking constraints. Put differently, they equal the marginal benefit (cost) to the central bank of not having to confirm private expectations that are essential to (1) and (2) when setting the current interest rate. When viewed in this way, the multipliers provide a signal of the temptation to renege on policy commitments. Large values of ψ_t occur during periods in which big reductions in loss are possible by succumbing to that temptation.

In this section I plot the sequence of conditional forecasts of ψ_t produced by the Kalman filter. The objective is to estimate the historical profile of the commitment multipliers and to interpret the observed variation.

Figure 7 graphs the forecasted series for the multipliers over the period 1982:Q1 to 2008:Q4. During the early 1980s the inflation multiplier (associated with the Phillips curve) was large, indicating that policymakers faced a strong incentive to abandon the commitment strategy by pushing interest rates higher in an effort to reduce inflation. The multiplier fell steadily throughout the mid 1980s, peaking a second time in 1990:Q3 and gradually declining thereafter until 1999:Q4. From 1993:Q4 to 2005:Q3 the inflation multiplier was negative. Because inflation was already low following a decade of tight policy and favorable macroeconomic conditions, the central bank could have benefited at the margin by deviating from the announced plan so as to raise inflation.

Excluding the mid to late 1990s, the contours of the output gap multiplier (associated with the IS equation) are broadly similar to those of the inflation multiplier. Dennis (2005) also finds evidence of positive correlation between the multipliers and traces this result to the central relationship modeled in the Phillips curve. All else constant, positive output gaps lead to higher inflation, so deviations from commitment aimed at increasing the output gap also tend to create more inflation. The same outcome could be reached by reneging on policy commitments with the goal of boosting inflation directly.

Appendix D. Reconciling Wald and LR Tests on λ_y

Likelihood ratio tests conducted in section 4.1 show that the data reject the null hypothesis of $\lambda_y = 0$ in the model with discretionary policy. This finding contradicts the outcome of an alternative test of parameter significance based on the Wald statistic. Formed by squaring the ratio of the point estimate of λ_y to its standard error, the Wald statistic is asymptotically distributed under the null as a chi-square random variable with one degree of freedom. The

p-value for a Wald test of $\lambda_y = 0$ is 0.602, indicating that the data fail to reject the null hypothesis at normal significance levels. The purpose of this section is to ascertain the source of the apparent contradiction between the Wald and likelihood ratio tests through a more careful examination of the log-likelihood function.

Table 8 reports estimates of the discretionary model for different restrictions on λ_y. During estimation I hold λ_y fixed at one of five distinct values while treating the other parameters as free. The values range from 0.0987 to zero, that is, from the unrestricted estimate to the lower bound of the parameter space. A better assessment of the statistical contribution of λ_y can be gained by comparing log likelihood across the columns of Table 8 corresponding to the various restrictions.

Setting $\lambda_y = 0.05$ appears to have little impact on parameter estimates relative to the unrestricted case. Possible exceptions include the interest rate smoothing coefficient $\lambda_{\Delta i}$ and the slope of the Phillips curve κ, both of which move towards zero as the weight on output gap stability declines. The p-value for a likelihood ratio test of $\lambda_y = 0.05$ is 0.806, revealing that the data has considerable difficulty distinguishing the unrestricted model from the null.

Once the restriction on λ_y reaches 0.02, the log-likelihood function begins to peak at two different locations in the parameter space. The high-density region settles on a point in which $(\lambda_{\Delta i}, \kappa) = (0.0136, 0.0009)$ and log likelihood equals -379.68. The low-density region settles at $(\lambda_{\Delta i}, \kappa) = (0.1470, 0.0092)$ with log likelihood equal to -383.15. The gap between high and low-density values for the other parameters is proportionally small. When high-density values are taken as the correct estimates of the restricted model, a test of the null hypothesis that $\lambda_y = 0.02$ cannot be rejected at standard significance levels. The p-value of the relevant likelihood ratio test is 0.709. By contrast, using low-density "estimates" would lead one to falsely conclude that the data reject the null with a p-value of 0.008. Similar results emerge for $\lambda_y = 0.005$. The global maximum is located at $(\lambda_{\Delta i}, \kappa) = (0.0035, 0.0002)$, where log likelihood is still not significantly different from its unrestricted value.

The last column of Table 8 shows that fixing $\lambda_y = 0$ in the high-density region causes $(\lambda_{\Delta i}, \kappa) \to (0, 0)$ during the course of estimation. Evaluating log likelihood at this point in the parameter space is problematic because the equilibrium of the model is not uniquely determined. To see why, note that setting $\kappa = 0$ reduces the Phillips curve (2) to a second-order difference equation of the form

$$\alpha E_t \pi_{t+1} - \pi_t + (1 - \alpha)\pi_{t-1} = -u_{\pi,t}. \tag{D.1}$$

Applying familiar techniques to solve (D.1) yields

$$\pi_t = \frac{1 - \alpha}{\alpha}\pi_{t-1} + \frac{1}{\alpha}u_{\pi,t}, \tag{D.2}$$

where $(1-\alpha)/\alpha$ is the stable root of the characteristic polynomial $\mathcal{P}(a) \equiv \alpha a^2 - a + (1 - \alpha)$. Meanwhile, setting $(\lambda_y, \lambda_{\Delta i}) = (0, 0)$ makes inflation the sole objective of policy as the loss function (3) simplifies to

$$\mathbb{L}_t = E_t(1 - \delta)\sum_{j=0}^{\infty}\delta^j \pi_{t+j}^2. \tag{D.3}$$

It is clear that minimizing (D.3) subject to (1) and (D.1) no longer determines a unique time path for the interest rate. Because inflation is now unaffected by changes in policy, the value of loss will be the same for any feasible sequence of interest rates. It follows from the IS equation (1) that any path for the output gap that satisfies

$$y_t + \sigma\tilde{i}_t = \phi E_t y_{t+1} + (1 - \phi)(\beta y_{t-1} + (1 - \beta)y_{t-2}) + \frac{\sigma(1-\alpha)^2}{\alpha^2}\pi_{t-1} + \frac{\sigma(1-\alpha)}{\alpha^2}u_{\pi,t} + u_{y,t} \tag{D.4}$$

is consistent with equilibrium, where $\tilde{i}_t$ represents any feasible realization of the interest rate.

Following most of the empirical new-Keynesian literature, I regard parameter combinations that generate multiple equilibria as strictly inadmissible for the purpose of estimation.

This is accomplished by simply truncating the log-likelihood function at the boundary of the determinacy region of the parameter space.[43] Consequently, the so-called high-density estimates for the case of $\lambda_y = 0$ are dismissed by construction. It turns out that the next highest point in the admissible parameter space occurs at the low-density estimates, where $(\lambda_{\Delta i}, \kappa) = (0.1767, 0.0111)$ and log likelihood equals -383.68.[44] Using low-density values to parameterize the restricted model leads to a p-value of 0.004 for the likelihood ratio test of $\lambda_y = 0$. These results shed light on the nature of the contradiction referenced at the beginning of the appendix. Ruling out estimates of structural parameters that do not satisfy equilibrium determinacy is what enables the likelihood ratio test to reject the hypothesis of zero weight on the output gap despite the large standard error associated with the unrestricted estimate of λ_y.

Appendix E. Computing the Posterior Density

Denote $\mathcal{L}(\theta|\mathcal{Y}^T)$ the likelihood function for a sample of observations $\mathcal{Y}^T = \{y_t^o, \pi_t^o, i_t^o\}_{t=1}^T$, and let $p(\theta)$ be the prior distribution over model parameters θ. According to Bayes' theorem, the posterior density $p(\theta|\mathcal{Y}^T)$ satisfies

$$p(\theta|\mathcal{Y}^T) \propto \mathcal{L}(\theta|\mathcal{Y}^T)p(\theta).$$

The random walk Metropolis-Hastings algorithm presented in An and Schorfheide (2007) is used to generate a sequence of draws from $p(\theta|\mathcal{Y}^T)$. The procedure works as follows.

[43] See Lubik and Schorfheide (2004) for an illustration of how the likelihood function can be extended to the indeterminacy region.

[44] As the restricted value of λ_y approaches zero, shifting $\lambda_{\Delta i}$ or κ away from their maximized values by small increments causes log likelihood to collapse. When $\lambda_y = $ 1e-5, for example, adjustments in $\lambda_{\Delta i}$ and κ on the order of 1e-6 reduce log likelihood from -379.71 to -1917.89. This points to a discontinuity in the likelihood function at the boundary of the determinacy region. The implication is that for all $\lambda_{\Delta i}$ and κ strictly greater than zero, the highest point on the likelihood surface corresponding to $\lambda_y = 0$ occurs at the low-density estimates $(\lambda_{\Delta i}, \kappa) = (0.1767, 0.0111)$.

1. Use a numerical optimization routine to maximize the log-posterior density kernel $\log \mathcal{L}(\theta|\mathcal{Y}^T) + \log p(\theta)$. The point that maximizes this function $\tilde{\theta}$ is the posterior mode.

2. Let $\tilde{\theta}$ be the starting value $\theta_{(0)}$ for the Markov chain whose stationary distribution converges to $p(\theta|\mathcal{Y}^T)$.

3. Draw $\theta^*_{(j)}$ from a proposal distribution $\mathcal{N}(\theta_{(j-1)}, c^2\Sigma)$, where the scalar c is chosen to keep the rejection rate of the algorithm between 50 and 80 percent. Setting $c = 0.10$ produces a rejection rate of about 60 percent in the model with an optimal simple rule.

4. Draw u from a uniform distribution with support $[0, 1]$.

5. Set $\theta_{(j)} = \theta^*_{(j)}$ if $u \leq r(\theta_{(j-1)}, \theta^*_{(j)}|\mathcal{Y}^T)$. Otherwise, set $\theta_{(j)} = \theta_{(j-1)}$. Here

$$r(\theta_{(j-1)}, \theta^*_{(j)}|\mathcal{Y}^T) = \min\left\{\frac{\mathcal{L}(\theta^*_{(j)}|\mathcal{Y}^T)p(\theta^*_{(j)})}{\mathcal{L}(\theta_{(j-1)}|\mathcal{Y}^T)p(\theta_{(j-1)})}, 1\right\}.$$

6. Repeat steps 3 through 5 for $j = 1, \ldots, N$, where N is the total number of draws in the Markov chain.

7. Approximate the posterior mean of a function $f(\theta)$ by $\frac{1}{N}\sum_{j=1}^{N} f(\theta_{(j)})$.

The covariance matrix Σ of the proposal distribution in step 3 is normally taken to be the inverse Hessian of the log-posterior density evaluated at the mode $\tilde{\theta}$. This assumption is not suitable for estimation under an optimal simple rule because, as explained in section 6.4, the likelihood function is not differentiable around the mode. Unfortunately, there is little guidance from the literature concerning alternative ways to parameterize the covariance matrix of the proposal distribution. As a result, I specify Σ as the element-by-element average of the inverse Hessian matrices obtained under commitment and discretion. The particular choice of Σ does not matter for estimation in the sense that the Metropolis-Hastings algorithm will produce, with a sufficient number of draws, a Markov chain whose ergodic distribution is $p(\theta|\mathcal{Y}^T)$. It does, however, affect the rate at which the empirical distribution converges to the ergodic distribution. As explained in Roberts, Gelman, and Gilks (1997), Σ can be fine-tuned (usually by adjusting c) to minimize the number of draws needed for the chain to converge.

I sampled three separate chains of 100,000 draws each, discarding the first 20,000 as burn-in, leaving 240,000 points from the posterior distribution. To assess convergence of the Markov chains, I conducted the partial means test developed by Geweke (1992) and the cumulative sum test proposed by Yu and Mykland (1998). I also considered informal graphical checks as suggested by An and Schorfheide (2007), namely, recursive means and pairwise scatter plots of the parameters for all three chains. Both diagnostic statistics and visual inspection of the graphs indicate convergence for all parameters.

References

Adam, Klaus and Billi, Roberto M. "Discretionary Monetary Policy and the Zero Lower Bound on Nominal Interest Rates." *Journal of Monetary Economics*, April 2007, 54(3), pp. 728-52.

Amato, Jeffrey D. and Laubach, Thomas. "Implications of Habit Formation for Optimal Monetary Policy." *Journal of Monetary Economics*, March 2004, 51(2), pp. 305-25.

An, Sungbae and Schorfheide, Frank. "Bayesian Analysis of DSGE Models." *Econometric Reviews*, March 2007, 26(2-4), pp. 113-72.

Ball, Laurence and Cecchetti, Stephen. "Inflation and Uncertainty at Long and Short Horizons." *Brookings Papers on Economic Activity 1*, 1990, 21, pp. 215-54.

Barro, Robert J. and Gordon, David B. "A Positive Theory of Monetary Policy in a Natural Rate Model." *Journal of Political Economy*, August 1983, 91(4), pp. 589-610.

Barsky, Robert B. "The Fisher Hypothesis and the Forecastibility and Persistence of Inflation." *Journal of Monetary Economics*, January 1987, 19(1), pp. 3-24.

Bernanke, Ben S.; Laubach, Thomas; Mishkin, Frederic S. and Posen, Adam S. *Inflation Targeting: Lessons from the International Experience*. Princeton, NJ: Princeton University Press, 1999.

Bernanke, Ben S. and Mishkin, Frederic S. "Inflation Targeting: A New Framework for Monetary Policy?" *The Journal of Economic Perspectives*, Spring 1997, 11(2), pp. 97-116.

Brainard, William C. "Uncertainty and the Effectiveness of Policy." *The American Economic Review*, May 1967, 57(2), pp. 411-25.

Brock, William A.; Durlauf, Steven N. and West, Kenneth D. "Policy Evaluation in Uncertain Economic Environments." *Brookings Papers on Economic Activity 1*, 2003, 34, pp. 235-322.

Calvo, Guillermo A. "Staggered Prices in a Utility-Maximizing Framework." *Journal of Monetary Economics*, September 1983, 12(3), pp. 383-98.

Castelnuovo, Efrem. "The Fed's Preference for Policy Rate Smoothing: Overestimation due to Misspecification?" *Topics in Macroeconomics*, 2006, 6(2), Article 5.

Castelnuovo, Efrem and Surico, Paolo. "Model Uncertainty, Optimal Monetary Policy and the Preferences of the Fed." *Scottish Journal of Political Economy*, February 2004, 51(1), pp. 105-26.

Cho, Seonghoon and Moreno, Antonio. "A Small-Sample Study of the New-Keynesian Macro Model." *Journal of Money, Credit, and Banking*, September 2006, 38(6), pp. 1461-81.

Christiano, Lawrence J.; Eichenbaum, Martin and Evans, Charles L. "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy." *Journal of Political Economy*, January 2005, 113(1), pp. 1-45.

Clarida, Richard; Galí, Jordi and Gertler, Mark. "The Science of Monetary Policy: A New Keynesian Perspective." *Journal of Economic Literature*, December 1999, 37(4), pp. 1661-707.

_________. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." *The Quarterly Journal of Economics*, February 2000, 115(1), pp. 147-80.

Culver, Sarah E. and Papell, David H. "Is There a Unit Root in the Inflation Rate? Evidence from Sequential Break and Panel Data Models." *Journal of Applied Econometrics*, July-August 1997, 12(4), pp. 435-44.

de Jong, Piet. "Smoothing and Interpolation with the State-Space Model." *Journal of the American Statistical Association*, December 1989, 84(408), pp. 1085-8.

Dennis, Richard. "Inferring Policy Objectives from Economic Outcomes." *Oxford Bulletin of Economics and Statistics*, September 2004, 66(S1), pp. 735-64.

_________. "Inflation Targeting under Commitment and Discretion." *Economic Review, Federal Reserve Bank of San Francisco*, 2005, pp. 1-13.

_________. "The Policy Preferences of the US Federal Reserve." *Journal of Applied Econometrics*, January 2006, 21(1), pp. 55-77.

_________. "When is Discretion Superior to Timeless Perspective Policymaking?" *Journal of Monetary Economics*, April 2010, 57(3), pp. 266-77.

Dennis, Richard and Söderström, Ulf. "How Important Is Precommitment for Monetary Policy?" *Journal of Money, Credit, and Banking*, June 2006, 38(4), pp. 847-72.

Dotsey, Michael. "Commitment Versus Discretion in Monetary Policy." *Business Review, Federal Reserve Bank of Philadelphia*, Fourth Quarter 2008, pp. 1-8.

Eggertsson, Gauti B. and Woodford, Michael. "The Zero Bound on Interest Rates and Optimal Monetary Policy." *Brookings Papers on Economic Activity 1*, 2003, 34, pp. 139-233.

Ehrmann, Michael and Smets, Frank. "Uncertain Potential Output: Implications for Monetary Policy." *Journal of Economic Dynamics and Control*, July 2003, 27(9), pp. 1611-38.

Estrella, Arturo and Fuhrer, Jeffrey C. “Dynamic Inconsistencies: Counterfactual Implications of a Class of Rational-Expectations Models.” *The American Economic Review*, September 2002, 92(4), pp. 1013-28.

Favero, Carlo A. and Rovelli, Riccardo. “Macroeconomic Stability and the Preferences of the Fed: A Formal Analysis, 1961-96.” *Journal of Money, Credit, and Banking*, August 2003, 35(4), pp. 545-56.

Fève, Patrick; Matheron, Julien and Sahuc, Jean-Guillaume. “Inflation Target Shocks and Monetary Policy Inertia in the Euro Area.” *The Economic Journal*, September 2010, 120(547), pp. 1100-24.

Fuhrer, Jeffrey C. “The (Un)Importance of Forward-Looking Behavior in Price Specifications.” *Journal of Money, Credit, and Banking*, August 1997, 29(3), pp. 338-50.

__________. “Habit Formation in Consumption and its Implications for Monetary-Policy Models.” *The American Economic Review*, June 2000, 90(3), pp. 367-90.

Fuhrer, Jeffrey C. and Moore, George R. “Inflation Persistence.” *The Quarterly Journal of Economics*, February 1995, 110(1), pp. 127-59.

Fuhrer, Jeffrey C. and Rudebusch, Glenn D. “Estimating the Euler Equation for Output.” *Journal of Monetary Economics*, September 2004, 51(6), pp. 1133-53.

Galí, Jordi and Gertler, Mark. “Inflation Dynamics: A Structural Econometric Analysis.” *Journal of Monetary Economics*, October 1999, 44(2), pp. 195-222.

Geweke, John F. “Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments,” in J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith, eds., *Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting*. Oxford and New York: Oxford University Press, 1992, pp. 169-94.

Giannoni, Marc P. and Woodford, Michael. “Optimal Inflation-Targeting Rules,” in Ben S. Bernanke and Michael Woodford, eds., *The Inflation Targeting Debate*. Chicago, IL: University of Chicago Press, 2005, pp. 93-162.

Givens, Gregory E. and Salemi, Michael K. “Generalized Method of Moments and Inverse Control.” *Journal of Economic Dynamics and Control*, October 2008, 32(10), pp. 3113-47.

Goodfriend, Marvin. “Inflation Targeting in the United States?” NBER Working Paper No. 9981, September 2003.

Hamilton, James D. *Time Series Analysis*. Princeton, NJ: Princeton University Press, 1994.

Ilbas, Pelin. "Estimation of Monetary Policy Preferences in a Forward-Looking Model: A Bayesian Approach." *International Journal of Central Banking*, September 2010a, 6(3), pp. 169-209.

_________. "Revealing the Preferences of the US Federal Reserve." forthcoming in *Journal of Applied Econometrics*, July 2010b.

Ingram, Beth Fisher; Kocherlakota, Narayana R. and Savin, N. E. "Explaining Business Cycles: A Multiple-Shock Approach." *Journal of Monetary Economics*, December 1994, 34(3), pp. 415-28.

Ireland, Peter N. "Changes in the Federal Reserve's Inflation Target: Causes and Consequences." *Journal of Money, Credit, and Banking*, December 2007, 39(8), pp. 1851-82.

Jensen, Henrik. "Targeting Nominal Income Growth or Inflation?" *The American Economic Review*, September 2002, 92(4), pp. 928-56.

Keen, Benjamin D. "Output, Inflation, and Interest Rates in an Estimated Optimizing Model of Monetary Policy." *Review of Economic Dynamics*, April 2009, 12(2), pp. 327-43.

Khan, Aubhik; King, Robert G. and Wolman, Alexander L. "Optimal Monetary Policy." *The Review of Economic Studies*, October 2003, 70(4), pp. 825-60.

Kiley, Michael T. "A Quantitative Comparison of Sticky-Price and Sticky-Information Models of Price Setting." *Journal of Money, Credit, and Banking*, February 2007, 39(S1), pp. 101-25.

Kimball, Miles S. "The Quantitative Analytics of the Basic Neomonetarist Model." *Journal of Money, Credit, and Banking*, November 1995, 27(4), pp. 1241-77.

Klein, Paul. "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model." *Journal of Economic Dynamics and Control*, September 2000, 24(10), pp. 1405-23.

Kydland, Finn E. and Prescott, Edward C. "Rules Rather than Discretion: The Inconsistency of Optimal Plans." *Journal of Political Economy*, June 1977, 85(3), pp. 473-91.

Levin, Andrew T.; Wieland, Volcker, and Williams, John C. "Robustness of Simple Monetary Policy Rules under Model Uncertainty," in John B. Taylor, ed., *Monetary Policy Rules*. Chicago and London: University of Chicago Press, 1999, pp. 263-299.

Levin, Andrew T. and Williams, John C. "Robust Monetary Policy with Competing Reference Models." *Journal of Monetary Economics*, July 2003, 50(5), pp. 945-75.

Lindé, Jesper. "Estimating New-Keynesian Phillips Curves: A Full Information Maximum Likelihood Approach." *Journal of Monetary Economics*, September 2005, 52(6), pp. 1135-49.

Lowe, Philip and Ellis, Luci. "The Smoothing of Official Interest Rates," in Philip Lowe, ed., *Monetary Policy and Inflation Targeting.* Reserve Bank of Australia, 1997, pp. 286-312.

Lubik, Thomas A. and Schorfheide, Frank. "Testing for Indeterminacy: An Application to U.S. Monetary Policy." *The American Economic Review*, March 2004, 94(1), pp. 190-217.

McCallum, Bennett T. "Issues in the Design of Monetary Policy Rules." in John B. Taylor and Michael Woodford, eds., *Handbook of Macroeconomics.* Amsterdam: Elsevier, 1999, Volume 1, Part 3, pp. 1483-1530.

McCallum, Bennett T. and Nelson, Edward. "An Optimizing IS-LM Specification for Monetary Policy and Business Cycle Analysis." *Journal of Money, Credit, and Banking*, August 1999, 31(3), pp. 296-316.

Mishkin, Frederic S. "Is the Fisher Effect for Real?: A Reexamination of the Relationship Between Inflation and Interest Rates." *Journal of Monetary Economics*, November 1992, 30(2), pp. 195-215.

__________. "Monetary Policy and the Dual Mandate." Speech at Bridgewater College, Bridgewater, Virginia, April 2007.

Onatski, Alexei and Williams, Noah. "Empirical and Policy Performance of a Forward-Looking Monetary Model." *Journal of Applied Econometrics*, January-February 2010, 25(1), pp. 145-76.

Orphanides, Athanasios. "Monetary Policy Evaluation with Noisy Information." *Journal of Monetary Economics*, April 2003, 50(3), pp. 605-31.

Ozlale, Umit. "Price Stability vs. Output Stability: Tales of Federal Reserve Administrations." *Journal of Economic Dynamics and Control*, July 2003, 27(9), pp. 1595-610.

Perron, Pierre. "Trends and Random Walks in Macroeconomic Time Series: Further Evidence from a New Approach." *Journal of Economic Dynamics and Control*, June-September 1988, 12(2-3), pp. 297-332.

Ravn, Morten; Schmitt-Grohé, Stephanie and Uribe, Martin. "Deep Habits." *The Review of Economic Studies*, January 2006, 73(1), pp. 195-218.

Roberts, G. O.; Gelman, A. and Gilks, W. R. "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms." *Annals of Applied Probability*, February 1997, 7(1), pp. 110-20.

Roberts, John M. "Is Inflation Sticky?" *Journal of Monetary Economics*, July 1997, 39(2), pp. 173-96.

_________. “How Well Does the New Keynesian Sticky-Price Model Fit the Data?” *Contributions to Macroeconomics*, 2005, 5(1), Article 10.

Rogoff, Kenneth. “The Optimal Degree of Commitment to an Intermediate Monetary Target.” *The Quarterly Journal of Economics*, November 1985, 100(4), pp. 1169-89.

Rotemberg, Julio J. and Woodford, Michael. “An Optimization-Based Econometric Framework for the Evaluation of Monetary Policy,” in Ben S. Bernanke and Julio J. Rotemberg, eds., *NBER Macroeconomics Annual.* Cambridge, MA: MIT Press, 1997, pp. 297-346.

Rudebusch, Glenn D. and Svensson, Lars E. O. “Policy Rules for Inflation Targeting,” in John B. Taylor, ed., *Monetary Policy Rules.* Chicago and London: University of Chicago Press, 1999, pp. 203-46.

Salemi, Michael K. “Econometric Policy Evaluation and Inverse Control,” *Journal of Money, Credit, and Banking*, October 2006, 38(7), pp. 1737-64.

Schaumburg, Ernst and Tambalotti, Andrea. “An Investigation of the Gains from Commitment in Monetary Policy.” *Journal of Monetary Economics*, March 2007, 54(2), pp. 302-24.

Smets, Frank and Wouters, Raf. “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area.” *Journal of the European Economic Association*, September 2003, 1(5), pp. 1123-75.

Söderlind, Paul. “Solution and Estimation of RE Macromodels With Optimal Policy.” *European Economic Review*, April 1999, 43(4-6), pp. 813-23.

Söderström, Ulf; Söderlind, Paul and Vredin, Anders. “New-Keynesian Models and Monetary Policy: A Reexamination of the Stylized Facts.” *The Scandinavian Journal of Economics*, September 2005, 107(3), pp. 521-46.

Steinsson, Jón. “Optimal Monetary Policy in an Economy with Inflation Persistence.” *Journal of Monetary Economics*, October 2003, 50(7), pp. 1425-56.

Stock, James H. and Watson, Mark W. “Testing for Common Trends.” *Journal of the American Statistical Association*, December 1988, 83(404), pp. 1097-107.

Svensson, Lars E. O. “Inflation Targeting as a Monetary Policy Rule.” *Journal of Monetary Economics*, June 1999, 43(3), pp. 607-54.

Taylor, John B. “Discretion Versus Policy Rules in Practice.” *Carnegie-Rochester Conference Series on Public Policy*, December 1993, 39(1), pp. 195-214.

Walsh, Carl E. "Speed Limit Policies: The Output Gap and Optimal Monetary Policy." *The American Economic Review*, March 2003, 93(1), pp. 265-78.

Woodford, Michael. "Optimal Monetary Policy Inertia." *The Manchester School*, 1999, 67(Supplement), pp. 1-35.

________. "Inflation Stabilization and Welfare." *Contributions to Macroeconomics*, 2002, 2(1), Article 1.

________. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton and Oxford: Princeton University Press, 2003a.

________. "Optimal Interest-Rate Smoothing." *The Review of Economic Studies*, October 2003b, 70(4), pp. 861-86.

Wynne, Mark A. "Core Inflation: A Review of Some Conceptual Issues." *Review, Federal Reserve Bank of St. Louis*, May-June 2008, 90(3-2), pp. 205-28.

Yu, B. and Mykland, P. "Looking at Markov Samplers Through CUSUM Path Plots: A Simple Diagnostic Idea." *Statistics and Computing*, August 1998, 8(3), pp. 275-86.

Yun, Tack. "Nominal Price Rigidity, Money Supply Endogeneity, and Business Cycles." *Journal of Monetary Economics*, April 1996, 37(2), pp. 345-70.

Table 1: Maximum-Likelihood Estimates

Parameter	Commitment		Discretion	
	Estimate	Std. Error	Estimate	Std. Error
σ_y	0.2482	0.0207	0.2065	0.0145
σ_π	0.5745	0.0395	0.6108	0.0505
$\sigma_{y\pi}$	-0.0381	0.0145	-0.0352	0.0162
σ_i	0.9274	0.0639	0.9403	0.0637
λ_y	0.1351	0.0596	0.0987	0.1892
$\lambda_{\Delta i}$	2.5581	0.5900	0.0579	0.0930
ϕ	0.3286	0.0102	0.3747	0.0071
β	1.5153	0.0101	1.4483	0.0083
σ	0.0089	0.0035	0.0002	0.0001
α	0.4927	0.0074	0.6162	0.0314
κ	0.0045	0.0016	0.0047	0.0090
$\log \mathcal{L}$	-387.7724		-379.6072	

Notes: The table reports maximum-likelihood estimates of the following model:

$$y_t = \phi E_t y_{t+1} + (1-\phi)(\beta y_{t-1} + (1-\beta) y_{t-2}) - \sigma(i_t - E_t \pi_{t+1}) + u_{y,t},$$

$$\pi_t = \alpha E_t \pi_{t+1} + (1-\alpha)\pi_{t-1} + \kappa y_t + u_{\pi,t},$$

$$\mathbb{L}_t = E_t(1-\delta)\sum_{j=0}^{\infty} \delta^j \{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2\}.$$

The term $\log \mathcal{L}$ denotes the maximized value of the log-likelihood function.

Table 2: Maximum-Likelihood Estimates of the Restricted Models

Parameter	Commitment	Discretion	Commitment	Discretion
σ_y	0.2607 (0.0292)	0.2024 (0.0137)	0.4509 (0.2485)	0.2063 (0.0144)
σ_π	0.5764 (0.0403)	0.6210 (0.0511)	0.5806 (0.0406)	0.6112 (0.0508)
$\sigma_{y\pi}$	-0.0423 (0.0157)	-0.0431 (0.0140)	-0.0664 (0.0484)	-0.0308 (0.0129)
σ_i	0.9423 (0.0654)	0.9582 (0.0654)	1.0149 (0.0708)	0.9385 (0.0633)
λ_y	0^*	0^*	0.0562 (0.0380)	0.0146 (0.0068)
$\lambda_{\Delta i}$	2.8536 (0.8119)	0.1767 (0.0741)	1^*	0.01^*
ϕ	0.3099 (0.0194)	0.3690 (0.0066)	0.1475 (0.2581)	0.3730 (0.0059)
β	1.5250 (0.0201)	1.4425 (0.0077)	1.6135 (0.1434)	1.4597 (0.0036)
σ	0.0067 (0.0038)	0.0002 (0.0001)	0.0561 (0.0516)	0.0003 (0.0001)
α	0.5010 (0.0068)	0.5906 (0.0267)	0.5106 (0.0076)	0.6250 (0.0261)
κ	0.0049 (0.0016)	0.0111 (0.0055)	0.0053 (0.0015)	0.0007 (0.0003)
$\log \mathcal{L}$	-389.8648	-383.6755	-396.5589	-379.6847

Notes: The table reports restricted maximum-likelihood estimates of the following model:

$$y_t = \phi E_t y_{t+1} + (1-\phi)(\beta y_{t-1} + (1-\beta) y_{t-2}) - \sigma(i_t - E_t \pi_{t+1}) + u_{y,t},$$

$$\pi_t = \alpha E_t \pi_{t+1} + (1-\alpha)\pi_{t-1} + \kappa y_t + u_{\pi,t},$$

$$\mathbb{L}_t = E_t(1-\delta) \sum_{j=0}^{\infty} \delta^j \{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2\}.$$

The superscript * denotes a parameter value that is imposed prior to estimation. The term $\log \mathcal{L}$ denotes the maximized value of the log-likelihood function. The numbers in parentheses are standard errors.

Table 3: Standard Deviations

Variable	Data	Commitment	Discretion
Inflation	1.0822	2.2140	1.2298
Output Gap	2.2506	4.5958	2.7194
Nominal Interest Rate	2.4144	3.7923	3.0239

Notes: Standard deviations are multiplied by 100.

Table 4: Model Comparison

Model	Log likelihood	BIC	Pseudo-odds ($z = 2$)	Pseudo-odds ($z = 3$)
Commitment	-387.7724	-413.4729	0.0003	0.0002
Discretion	-379.6072	-405.3078	0.9997	0.7489
Unrestricted	-371.3551	-406.4013	—	0.2509

Notes: BIC refers to the Bayesian information criterion. Larger values of BIC imply greater fit with the data. The pseudo-odds statistic measures the data-determined probability of a model j and is defined as $\rho(j) = \exp(BIC(j))/\sum_{h=1}^{z}\exp(BIC(h))$, where z is the number of distinct models under consideration.

Table 5: Counterfactual Losses under Discretion

Policy	Var(π)	Var(y)	Var(Δi)	Var(i)	Loss	π^{eq}	y^{eq}
No Zero Lower Bound							
Discretion	1.5124	7.3958	2.3421	9.1447	2.2148	0.4392	1.3978
Commitment	1.3886	6.7863	2.4175	66.7330	2.0219	—	—
Zero Lower Bound Imposed							
Discretion	1.5124	7.3958	2.3421	9.1447	2.2148	0.3665	1.1665
Commitment	1.4045	7.2425	1.8775	14.5811	2.0805	—	—

Notes: The table reports the variances of inflation π, the output gap y, and the interest rate in first differences Δi and levels i as well as the value of loss under discretion and commitment. The best policy is the one that produces the smallest value of loss. The loss differential is also reported in terms of an inflation equivalent π^{eq} and an output gap equivalent y^{eq}. The top panel contains results for the case with no zero bound constraint. The bottom panel contains results for the case in which the zero bound restriction is imposed by amending the loss function to include a suitable weight on the variance of the interest rate.

Table 6: Maximum-Likelihood Estimates of Alternative Models

	CPI Data		PCE Data		Augmented Loss	
Parameter	Commitment	Discretion	Commitment	Discretion	Commitment	Discretion
σ_y	0.2371 (0.0175)	0.2064 (0.0145)	0.2286 (0.0162)	0.2111 (0.0147)	0.2482 (0.0207)	0.2061 (0.0177)
σ_π	1.2982 (0.0904)	1.4353 (0.1417)	0.4683 (0.0388)	0.4278 (0.0326)	0.5745 (0.0395)	0.6125 (0.0568)
$\sigma_{y\pi}$	0.0145 (0.0304)	−0.0108 (0.0333)	−0.0171 (0.0108)	−0.0097 (0.0111)	−0.0381 (0.0145)	−0.0402 (0.0190)
σ_i	1.0413 (0.0715)	1.0796 (0.0735)	0.7975 (0.0544)	0.8905 (0.0609)	0.9274 (0.0639)	0.9426 (0.0770)
λ_y	0.0287 (0.1144)	0.0524 (0.1865)	0.3285 (0.1312)	0.3566 (0.3933)	0.1351 (0.0596)	0.1770 (0.2245)
$\lambda_{\Delta i}$	4.2837 (0.9977)	0.1789 (0.0748)	11.6085 (0.4593)	0.7848 (0.9047)	2.5581 (0.5900)	1.23e-5 (0.1967)
λ_i	0^*	0^*	0^*	0^*	$0^\dagger$	1.30e-7 (0.0021)
ϕ	0.3344 (0.0051)	0.3764 (0.0068)	0.3418 (0.0044)	0.3725 (0.0072)	0.3286 (0.0102)	0.3772 (0.0067)
β	1.5393 (0.0114)	1.4413 (0.0092)	1.5334 (0.0119)	1.4563 (0.0074)	1.5153 (0.0101)	1.4409 (0.0115)
σ	0.0093 (0.0029)	0.0002 (0.0001)	0.0081 (0.0025)	0.0017 (0.0006)	0.0089 (0.0035)	2.08e-8 (0.0003)
α	0.4996 (0.0081)	0.6693 (0.0477)	0.4313 (0.0277)	0.5063 (0.0299)	0.4927 (0.0074)	0.6170 (0.0556)
κ	0.0129 (0.0045)	0.0282 (0.0140)	0.0127 (0.0065)	0.0060 (0.0078)	0.0045 (0.0016)	0.0080 (0.0087)
$\log \mathcal{L}$	−487.8680	−479.9667	−347.6384	−353.3437	−387.7724	−379.1178
BIC	−513.5686	−505.6673	−373.3390	−379.0443	−413.4729	−407.1548
ρ	0.0004	0.9996	0.9967	0.0033	0.0018	0.9982

Notes: The table reports maximum-likelihood estimates of the following model:

$$y_t = \phi E_t y_{t+1} + (1-\phi)(\beta y_{t-1} + (1-\beta) y_{t-2}) - \sigma(i_t - E_t \pi_{t+1}) + u_{y,t},$$

$$\pi_t = \alpha E_t \pi_{t+1} + (1-\alpha)\pi_{t-1} + \kappa y_t + u_{\pi,t},$$

$$\mathbb{L}_t = E_t(1-\delta)\sum_{j=0}^{\infty} \delta^j \{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2 + \lambda_i i_{t+j}^2\}.$$

The superscript * denotes a parameter value that is imposed prior to estimation. The superscript † denotes a value that converged to the boundary of the allowable parameter space during estimation. The numbers in parentheses are standard errors. The term $\log \mathcal{L}$ denotes the maximized value of the log-likelihood function. BIC refers to the Bayesian information criterion. The pseudo-odds statistic measures the data-determined probability of a model j and is defined as $\rho(j) = \exp(BIC(j))/\sum_{h=1}^{z} \exp(BIC(h))$, where z is the number of distinct models under consideration.

Table 7: Estimation Results for an Optimal Simple Rule

Parameter	Prior Dist.	Post. Mean	95% Prob. Int.	Post. Mode	Commitment	Discretion
σ_y	$[0.10, 0.90]$	0.4548	$[0.34, 0.58]$	0.4292	0.2482 (0.0207)	0.2065 (0.0145)
σ_π	$[0.10, 0.90]$	0.5757	$[0.51, 0.64]$	0.5637	0.5745 (0.0395)	0.6108 (0.0505)
$\sigma_{y\pi}$	$[-0.20, 0.00]$	-0.0646	$[-0.12, -0.02]$	-0.0675	-0.0381 (0.0145)	-0.0352 (0.0162)
σ_i	$[0.50, 1.50]$	1.0443	$[0.93, 1.16]$	1.0249	0.9274 (0.0639)	0.9403 (0.0637)
λ_y	$[0.01, 0.20]$	0.0964	$[0.06, 0.14]$	0.0960	0.1351 (0.0596)	0.0987 (0.1892)
$\lambda_{\Delta i}$	$[0.25, 1.25]$	0.7485	$[0.71, 0.77]$	0.7599	2.5581 (0.5900)	0.0579 (0.0930)
ϕ	$[0.01, 0.40]$	0.1259	$[0.02, 0.26]$	0.1389	0.3286 (0.0102)	0.3747 (0.0071)
β	$[1.30, 1.70]$	1.5024	$[1.37, 1.63]$	1.5025	1.5153 (0.0101)	1.4483 (0.0083)
σ	$[0.01, 0.10]$	0.0492	$[0.03, 0.07]$	0.0460	0.0089 (0.0035)	0.0002 (0.0001)
α	$[0.30, 0.70]$	0.5132	$[0.50, 0.53]$	0.5109	0.4927 (0.0074)	0.6162 (0.0314)
κ	$[0.001, 0.010]$	0.0020	$[0.001, 0.004]$	0.0016	0.0045 (0.0016)	0.0047 (0.0090)
$\log \mathcal{L}$				-395.6591	-387.7724	-379.6072
BIC				-421.3596	-413.4729	-405.3078
ρ				0.0000	0.0003	0.9997

Notes: The table reports estimation results for the following model:

$$y_t = \phi E_t y_{t+1} + (1-\phi)(\beta y_{t-1} + (1-\beta) y_{t-2}) - \sigma(i_t - E_t \pi_{t+1}) + u_{y,t},$$

$$\pi_t = \alpha E_t \pi_{t+1} + (1-\alpha)\pi_{t-1} + \kappa y_t + u_{\pi,t},$$

$$i_t = \theta_i i_{t-1} + (1-\theta_i)(\theta_\pi \pi_{t-1} + \theta_{y1} y_{t-1} + \theta_{y2} y_{t-2}),$$

$$\mathbb{L}_t = E_t(1-\delta)\sum_{j=0}^{\infty} \delta^j \{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2\},$$

where $\{\theta_i, \theta_\pi, \theta_{y1}, \theta_{y2}\} = \arg\min \mathbb{L}_t$. The first column displays the support of the uniform prior distribution for each parameter. The next two columns report the posterior means and 95 percent probability intervals generated from the random walk Metropolis-Hastings algorithm described in An and Schorfheide (2007). The fourth column presents the maximum-likelihood estimates, which are equivalent to the posterior mode estimates under the assumption of independent uniform priors. The last two columns reproduce the estimates under full commitment and discretion, where the numbers in parentheses correspond to the Hessian-based standard errors. The term $\log \mathcal{L}$ denotes the maximized value of the log-likelihood function. BIC refers to the Bayesian information criterion. The pseudo-odds statistic measures the data-determined probability of a model j and is defined as $\rho(j) = \exp(BIC(j))/\sum_{h=1}^{z}\exp(BIC(h))$, where z is the number of distinct models under consideration.

Table 8: Restricted Maximum-Likelihood Estimates

Parameter	$\lambda_y = 0.0987$	$\lambda_y = 0.05$	$\lambda_y = 0.02$		$\lambda_y = 0.005$		$\lambda_y = 0$	
σ_y	0.2065	0.2063	0.2063	0.2025	0.2063	0.2024	0.2063	0.2024
σ_π	0.6108	0.6111	0.6112	0.6202	0.6112	0.6208	0.6113	0.6210
$\sigma_{y\pi}$	−0.0352	−0.0327	−0.0311	−0.0408	−0.0303	−0.0427	−0.0300	−0.0431
σ_i	0.9403	0.9391	0.9386	0.9561	0.9384	0.9578	0.9383	0.9582
$\boldsymbol{\lambda_{\Delta i}}$	**0.0579**	**0.0321**	**0.0136**	**0.1470**	**0.0035**	**0.1707**	**$0^\dagger$**	**0.1767**
ϕ	0.3747	0.3736	0.3730	0.3673	0.3728	0.3687	0.3727	0.3690
β	1.4483	1.4503	1.4515	1.4451	1.4521	1.4429	1.4523	1.4425
σ	0.0002	0.0002	0.0003	0.0002	0.0003	0.0002	0.0003	0.0002
α	0.6162	0.6213	0.6244	0.5991	0.6259	0.5922	0.6265	0.5906
$\boldsymbol{\kappa}$	**0.0047**	**0.0024**	**0.0009**	**0.0092**	**0.0002**	**0.0108**	**$0^\dagger$**	**0.0111**
$\log \mathcal{L}$	−379.6072	−379.6374	−379.6770	−383.1542	−379.6984	−383.5578	−379.7053	−383.6755

Notes: The table reports restricted maximum-likelihood estimates of the following model:

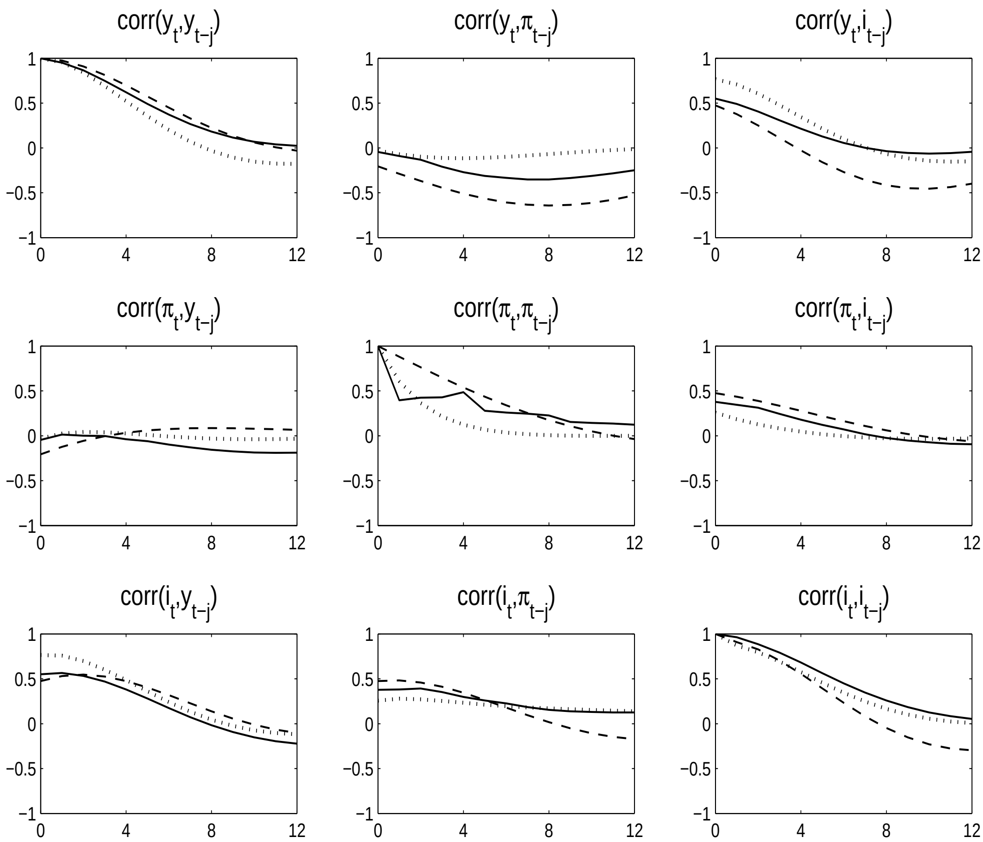
$$y_t = \phi E_t y_{t+1} + (1-\phi)(\beta y_{t-1} + (1-\beta)y_{t-2}) - \sigma(i_t - E_t\pi_{t+1}) + u_{y,t},$$

$$\pi_t = \alpha E_t \pi_{t+1} + (1-\alpha)\pi_{t-1} + \kappa y_t + u_{\pi,t},$$

$$\mathbb{L}_t = E_t(1-\delta)\sum_{j=0}^{\infty}\delta^j\{\pi_{t+j}^2 + \lambda_y y_{t+j}^2 + \lambda_{\Delta i}(i_{t+j} - i_{t+j-1})^2\},$$

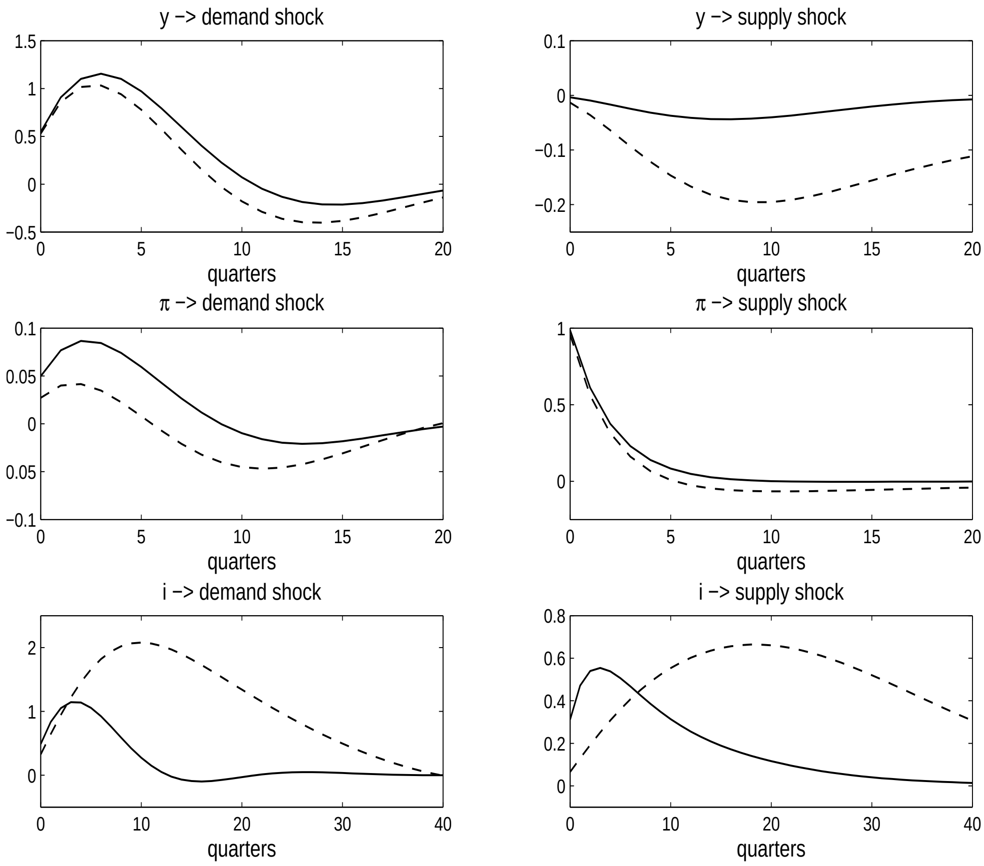
where $\{i_{t+j}\}_{j=0}^{\infty}$ minimizes $\mathbb{L}_t$ under discretion. The superscript † denotes a value that converged to the boundary of the allowable parameter space during estimation. The term $\log \mathcal{L}$ denotes the maximized value of log likelihood. Double columns of estimates corresponding to $\lambda_y = \{0.02, 0.005, 0\}$ indicate that the likelihood function peaks at two different points when these particular restrictions are imposed.

Figure 1: Vector Autocorrelation Functions



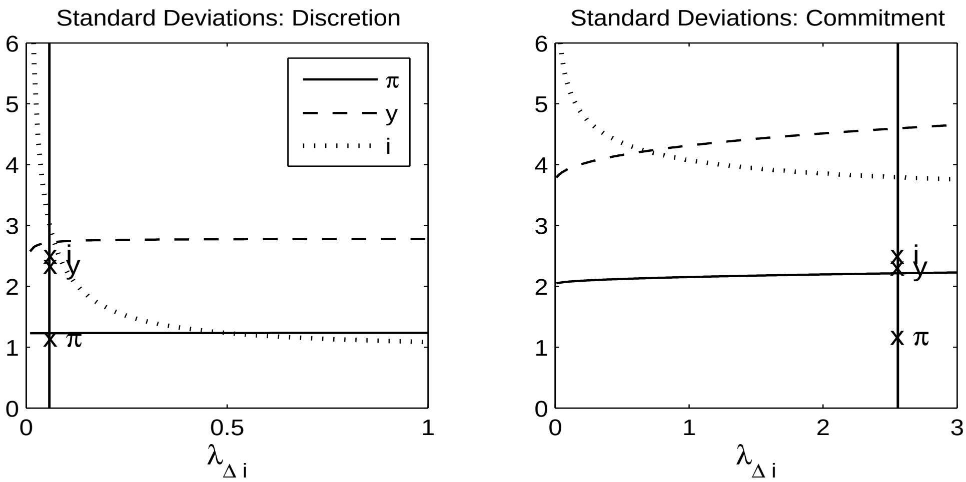
Notes: The figure shows the vector autocorrelation functions for the output gap y, inflation π, and the nominal interest rate i implied by the US data (solid line), the estimated model under commitment (dashed line), and the estimated model under discretion (dotted line).

Figure 2: Impulse Response Functions



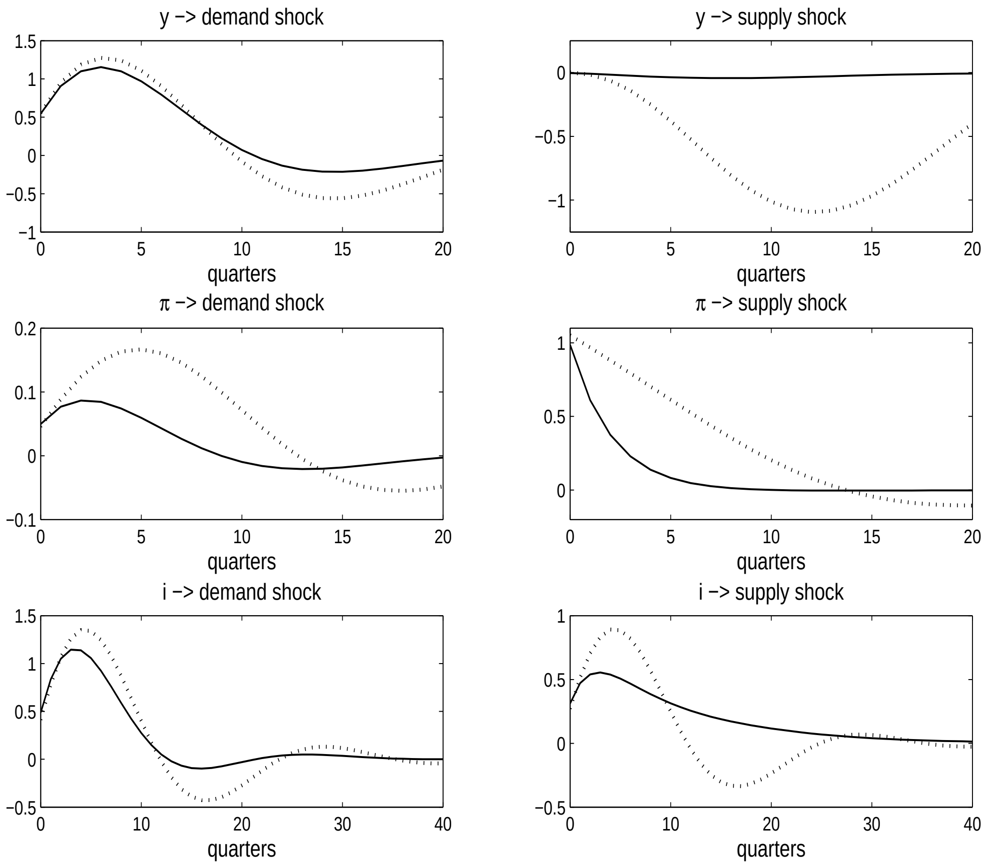
Notes: The figure displays the impulse responses of the output gap y, inflation π, and the nominal interest rate i to a demand shock $u_{y,t}$ (left column) and a supply shock $u_{\pi,t}$ (right column). Response functions are graphed for the estimated model under discretion (solid line) and the commitment model using the same point estimates obtained under discretion (dashed line). Each panel traces out the effect of an estimated one standard deviation shock, and the values are interpreted as percent deviations from the steady state.

Figure 3: Varying the Weight on Interest Rate Smoothing



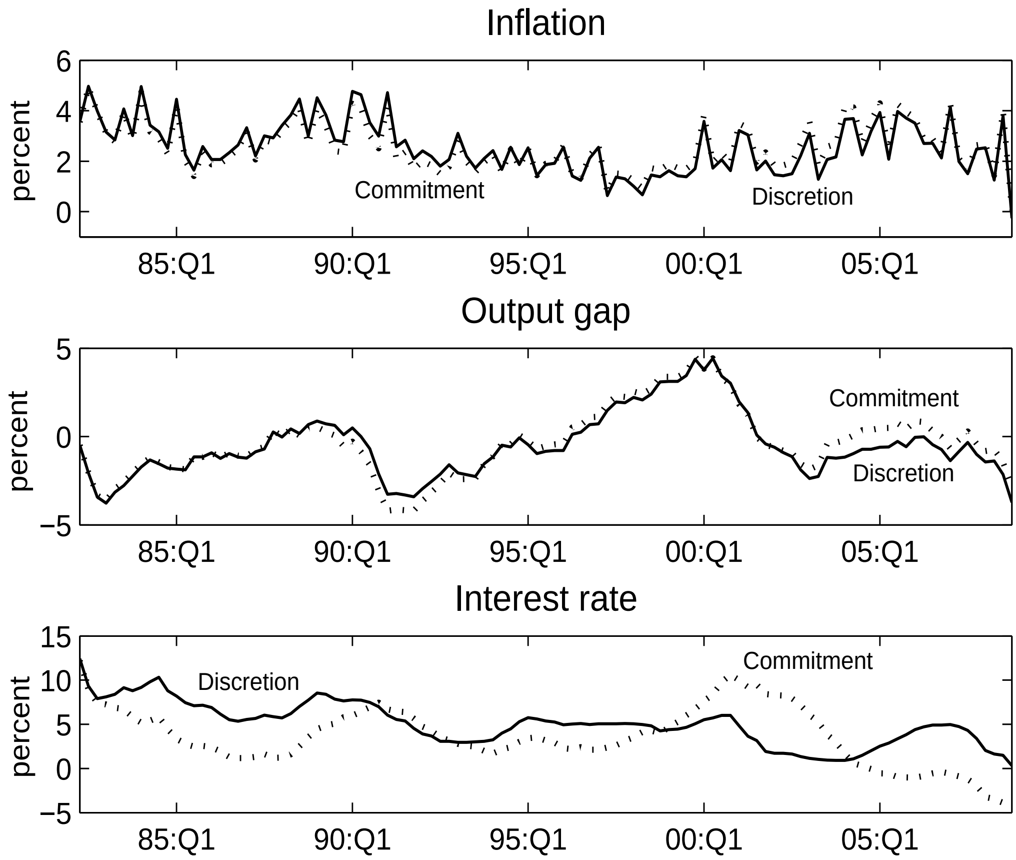
Notes: The figure displays the standard deviations of inflation (solid line), the output gap (dashed line), and the nominal interest rate (dotted line) as the weight on interest rate smoothing $\lambda_{\Delta i}$ varies from its estimated value, holding the other parameters fixed at their point estimates. The left panel graphs the functions implied by the estimated model under discretion. The right panel corresponds to the commitment model. Vertical lines indicate the estimated values of $\lambda_{\Delta i}$ and the crosses identify sample moments.

Figure 4: Impulse Response Functions



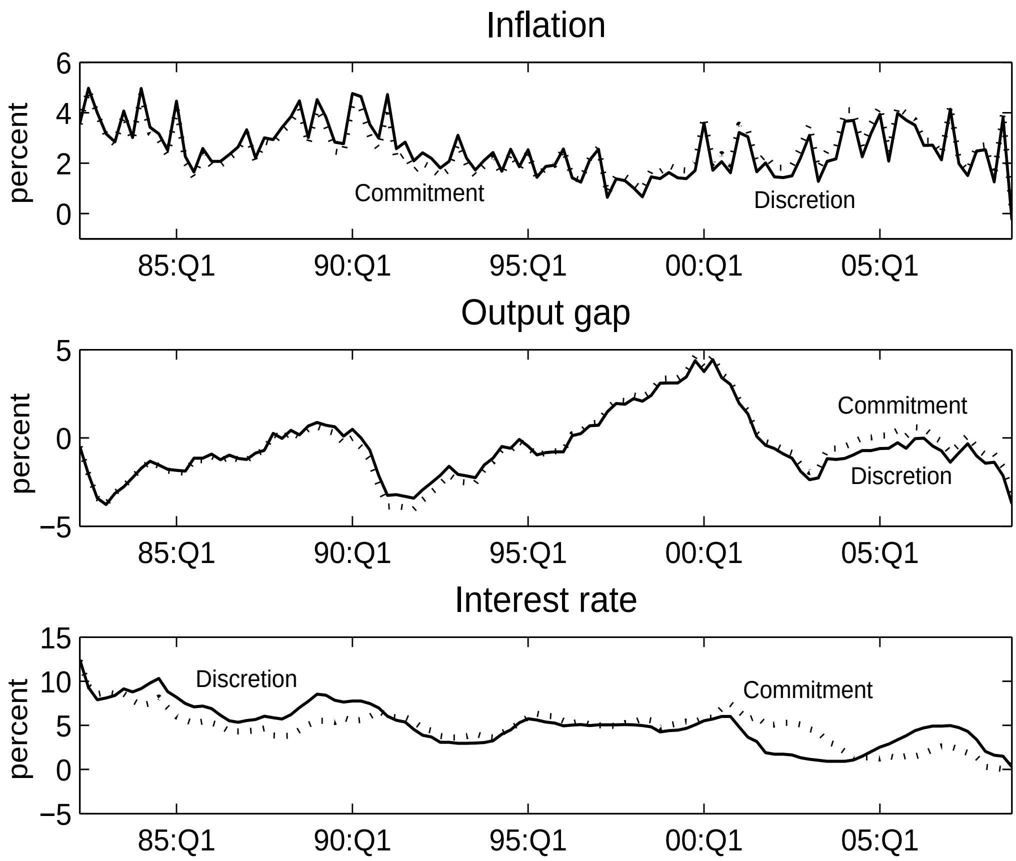
Notes: The figure displays the impulse responses of the output gap y, inflation π, and the nominal interest rate i to a demand shock $u_{y,t}$ (left column) and a supply shock $u_{\pi,t}$ (right column). Response functions are graphed for the estimated model under discretion (solid line) and the estimated model under commitment (dotted line). Each panel traces out the effect of an estimated one standard deviation shock, and the values are interpreted as percent deviations from the steady state.

Figure 5: Counterfactual Simulations



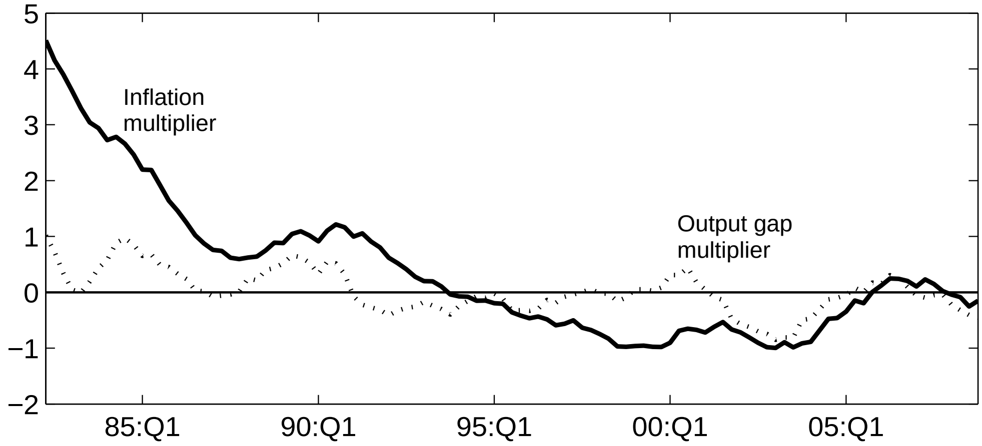
Notes: The figure plots the actual series as implied by the estimated discretionary model (solid lines) and the counterfactual series generated under commitment (dotted lines). Counterfactual simulations are obtained by estimating the historical shocks in the discretionary model using the fixed interval Kalman smoother. The shocks are then reinserted into the model but with the central bank minimizing loss under commitment.

Figure 6: Counterfactual Simulations with a Zero Bound Constraint



Notes: The figure plots the actual series as implied by the estimated discretionary model (solid lines) and the counterfactual series generated under commitment (dotted lines) with the zero lower-bound restriction. Counterfactual simulations are obtained by estimating the historical shocks in the discretionary model using the fixed interval Kalman smoother. The shocks are then reinserted into the model but with the central bank minimizing loss under commitment. The loss function is amended to include a weight on the variance of the interest rate just large enough to ensure that the simulated path never violates the zero bound.

Figure 7: Commitment Multipliers



Notes: The figure plots conditional forecasts of the commitment multipliers associated with the IS equation (dotted line) and the Phillips curve (solid line). The sequence of forecasts are computed recursively from the Kalman filter using the commitment model estimated in Table 1.