Generalized method of moments and inverse control

Gregory E. Givens[a], Michael K. Salemi[b,*]

[a] *Department of Economics and Finance, Middle Tennessee State University*

[b] *Department of Economics, University of North Carolina, Chapel Hill, NC 27599-3305, USA*

Abstract

This paper presents a Generalized Method of Moments algorithm for estimating the structural parameters of a macroeconomic model subject to the restriction that the coefficients of the monetary policy rule minimize the central bank's expected loss function. The algorithm combines least-squares normal equations with moment restrictions derived from the first-order necessary conditions of the auxiliary optimization. We assess the performance of the algorithm with Monte Carlo simulations using three increasingly complex models. We find that imposing the optimizing restrictions when they are true improves estimation accuracy and that imposing those restrictions when they are false biases estimates of some of the structural parameters but not of the policy-rule coefficients.

JEL Classification: C32; C61; E31; E32; E52; E61

Keywords: GMM; optimal monetary policy; simple rules; policy objectives

1. Introduction

Two of the most influential developments in the field of monetary economics have been the rise of interest rate rules as a means of capturing the systematic component of policy (e.g., Taylor (1993) and Clarida, Galí, and Gertler (2000)) as well as the advancement of structural New Keynesian models usable for policy analysis (e.g., Fuhrer and Moore (1995) and Rotemberg and Woodford (1997)). A recurring theme emerging from this literature is that the historical conduct of monetary policy in the U.S., as reflected by the estimated coefficients of an interest rate rule, is very different from the behavior implied by optimal

*Corresponding author. Tel.: + 919 966 5391; fax: + 919 966 4986.
Email address: Michael_Salemi@unc.edu (M.K. Salemi).

rules derived in a framework that couples an empirical model of the economy with an explicit loss function for the central bank. Rudebusch (2001) notes that optimized rules normally call for an aggressive response to departures of inflation and output from target. Estimated rules, on the other hand, indicate not only a more conservative reaction to both variables, but also a tendency to avoid volatile swings in the policy instrument, a phenomenon known as interest rate smoothing.

Dennis (2005) argues that the reason for the apparent disconnect between optimal and historical policies is that counterfactual policy analysis is often carried out using a parameterized loss function that is inconsistent with outcomes observed in U.S. data. Attempts to reconcile estimated rules with ones that solve an optimal policy exercise have produced a burgeoning literature where the weights in the central bank's objective function are chosen with an eye to the data. Favero and Rovelli (2003), Ozlale (2003), and Dennis (2005) estimate backward-looking models of aggregate demand and supply subject to an auxiliary condition that the policy rule minimizes a quadratic loss function. Imposing an optimality restriction in the course of estimation enables them to obtain joint estimates of the structural parameters and the policy weights that identify central bank preferences. The result is an empirical model that is jointly consistent with the historical record and policy optimality.

Salemi (2006) and Dennis (2004) demonstrate how to generalize the estimation procedure to incorporate forward-looking models embodying rational expectations. Although there are obvious advantages to modeling policy decisions at a structural level, this simultaneous approach to estimation comes with a cost. The task of nesting the optimal-policy algorithm of the central bank within the estimation algorithm of the researcher is computationally burdensome. One commonly-used strategy involves estimating the parameters of the model (including the loss function weights) by quasi-maximum likelihood, in which case the optimal policy-rule coefficients are found each time a sample likelihood value is recorded. This "brute force" approach is computationally intensive because the great majority of estimation time involves identifying optimal policies for parameter values that do not fit the data.

Since Hansen (1982) refocused attention on method of moments estimation, the Generalized Method of Moments (GMM) has become an important component of the econometrician's toolkit. In this paper, we develop a new GMM algorithm for estimating the structural parameters of three increasingly complex New Keynesian-style models while maintaining the strict assumption that policy-rule coefficients minimize the central bank's expected loss function. The algorithm combines the least squares normal equations implied by the model's reduced form with the first-order necessary conditions characterizing the policymaker's optimal choice of coefficients. The result is a parsimonious set of orthogonality conditions that form the basis for estimation using GMM.

The empirical strategy adopted here is computationally more efficient than the brute force approach because it circumvents the need to perform an optimal control exercise for each set of parameters considered in the course of estimation. Instead, the algorithm searches freely over values of the structural parameters, the loss function weights, and the policy-rule coefficients for those that satisfy a collection of moment conditions, a subset of which correspond to the first order conditions of the policymaker's control problem. Although the algorithm bypasses an explicit optimal control exercise, it is still an example of *inverse control* because it permits the researcher to obtain estimates of monetary policy objectives by observing the actions embodied in the policy rule.[1]

Our estimation algorithm is perhaps most similar to the method expounded by Favero and Rovelli (2003). They use GMM to estimate an aggregate demand and supply model featuring an interest-rate targeting criterion derived from minimization of the central bank's loss function. Our contribution departs from their's along two critical dimensions. First, their algorithm can only be applied to backward-looking models of the economy. In contrast, our

[1]Refer to Salemi (2006) for a more detailed discussion of inverse control theory and its applications in monetary policy analysis.

approach generalizes to a broader class of models that include backward and forward-looking structures. Second, to obtain an estimable set of orthogonality conditions, the authors assume that central bank preferences are defined over an intertemporal loss function with an arbitrary finite horizon. The estimation algorithm discussed in this paper, on the other hand, can be used to identify the policy weights of a more conventional objective function defined over an infinite horizon.

By means of Monte Carlo simulations, we assess the performance of GMM in two cases. In the first case, the true policy coefficients are optimal for a given set of loss function weights. In the second, the true coefficients are not optimal for any policy weights. The second case is important because it demonstrates what can happen if one conditions estimation on the erroneous assumption that observed policy actions are the outcome of optimal behavior. We find that when the hypothesis of policy optimality is true, GMM consistently returns unbiased estimates of all parameters, with each converging to its true value as sample size increases. We also find that imposing policy optimality sharpens estimates of several structural parameters. In contrast, imposing optimality when it is false yields biased estimates of some parameters but does not lead to bias of the policy-rule coefficients.

Finally, we demonstrate the breadth of our GMM methodology with an application to actual U.S. data. Such an exercise allows us to clarify the practical significance of including optimal policy restrictions by comparing outcomes in two cases. In the first case, policy-rule coefficients are unrestricted, and in the second, they are constrained to satisfy the first order conditions for a loss-minimizing policy. For each model considered, we find that imposing optimality conditions weakens empirical performance according to standard measures of model fit. We also find that policy optimality changes the estimates of some key structural parameters and policy-rule coefficients in a meaningful way.

The remainder of the paper is organized as follows. Section 2 describes the econometric problem and demonstrates how to write the policymaker's first order conditions as moment restrictions. Section 3 documents Monte Carlo results obtained after applying our GMM algorithm to three different models of the New Keynesian variety. Section 4 reports estimates for each model when taken to U.S. data. Section 5 compares our GMM strategy with two alternative procedures commonly used in the literature, namely, full information maximum likelihood and a two-step estimation approach. Section 6 concludes.

2. The econometric problem

We consider a macroeconomic model characterized by a system of dynamic, discrete-time rational expectations equations of the form:

$$E_t F\left(X_{t-1}, X_t, X_{t+1}, \ldots, X_{t+j}, \varepsilon_t; \rho, \theta\right) = 0 \tag{1}$$

where X_t and ε_t are $(n_x \times 1)$ and $(n_\varepsilon \times 1)$ vectors of endogenous variables and serially uncorrelated exogenous shocks, respectively. Let E_t denote the mathematical expectations operator conditional on information available through date t, and let $j > 1$ be the maximum number of leads necessary to describe the n_f equations stacked in the vector F. The secondary arguments, ρ and θ, are vectors whose elements correspond to the underlying structural parameters and the coefficients of the monetary policy rule, respectively. Given initial conditions and a sequence of exogenous shocks $\{\varepsilon_t\}_{t=1}^{\infty}$, equation (1) determines $\{X_t\}_{t=1}^{\infty}$.

One equation in F describes the behavior of the central bank and is given by the following rule for setting the nominal interest rate:

$$r_t = P\left(X_{t-1}, w_t; \theta\right) \tag{2}$$

The elements of θ govern how the policymaker adjusts the interest rate to changes in economic events as represented by X_{t-1}, and w_t is a white-noise disturbance summarizing all exogenous variation in the policy instrument. Throughout the paper we maintain the assumption that the central bank commits to a simple, fixed-parameter rule and chooses θ optimally to achieve an explicit objective that we describe later.

We justify the decision to represent monetary policy using a simple instrument rule on numerous grounds. First, Levin, Wieland, and Williams (1999) argue that simple policy rules incorporating feedback from a small set of variables perform well across a variety of macroeconomic models featuring rational expectations. The implication is that simple rules are more robust to uncertainty regarding the true structure of the economy. Second, as noted by McCallum (1999), simple rules of the form (2) are operational in the sense that they identify the policy instrument r_t as a variable that the central bank can actually control, and by way of feedback from X_{t-1}, require only information about the state of the economy that is readily observable at the beginning of period t. Finally, by construction, simple rules have the desirable property that they can be communicated to the public and verified without much difficulty, enhancing visibility of central bank actions.

For the models that we consider, the rational expectations solution to (1) can be expressed as a first-order autoregression

$$X_t = GX_{t-1} + H\varepsilon_t \tag{3}$$

where G and H are $(n_x \times n_x)$ and $(n_x \times n_\varepsilon)$ matrices of reduced-form parameters whose elements are nonlinear functions of ρ and θ. For the purpose of estimation, we denote the residual term $\varphi_t \equiv H\varepsilon_t$ as the $(n_x \times 1)$ vector of reduced-form errors with covariance matrix Φ. Because φ_t is a vector containing unique linear combinations of the elements of ε_t, it is serially uncorrelated.

In the spirit of Svensson (1999) and Clarida, Galí, and Gertler (1999), we assume that the stabilization objectives of the central bank are summarized by the following quadratic loss function:

$$\Lambda = E_t \sum_{j=0}^{\infty} \delta^j \left(X_{t+j} - X^*\right)' W \left(X_{t+j} - X^*\right) \tag{4}$$

where X^* is a vector of fixed target values for X_{t+j} and $\delta \in (0, 1)$ is a discount factor. W is a $(n_x \times n_x)$ matrix of loss function coefficients whose elements contain the nonnegative policy weights that measure the relative importance of each objective. Optimal policy means searching over the elements of θ for those that minimize Λ and guarantee a unique rational expectations equilibrium to (1). The resulting policy is not the unconstrained optimal commitment policy in a global sense, but rather the best policy that resides within a family of simple instrument rules of the form (2).

Because we are interested in retrieving the loss function weights implied by the estimate of θ, we must take a stand on how policy is conducted in our sample, recognizing that the nature of policy determines the true data generating process. There are generally three classes of policies that have received considerable attention in the literature. The first is a once-and-for-all commitment policy in which the central bank chooses a state-contingent path $\{X_t\}_{t=0}^{\infty}$ subject to the constraints implied by the structural model. Svensson and Woodford (2005) show that such a policy can be represented by an explicit instrument rule that responds to not only the current state vector, but the entire history of state vectors dating back to the announcement of the policy. The second class, optimal discretion, requires the central bank to choose the interest rate by reoptimizing its loss function every period taking as given private sector expectations of future variables (e.g., Jensen (2002), Walsh (2003), and Vestin (2006)). The idea is that past policy decisions in no way constrain current or future policy decisions like they do under commitment. The third class of policies can be broadly defined as commitment to a simple, fixed-parameter instrument rule along the lines of Taylor (1993). While retaining many of the expectations-forming benefits of optimal commitment, simple interest rate rules involve feedback from only a small set of observable variables (e.g., McCallum and Nelson (1999b) and Schmitt-Grohé and Uribe (2007)).

In this paper, we assume that actual policy is conducted according to a simple instrument rule of the form (2). That is, given the loss function (4), the true data generating process is determined by commitment to a simply policy rule whose coefficients are chosen to minimize expected loss. If instead we assumed that policy is the outcome of optimal commitment (e.g., Ilbas (2007)) or optimal discretion (e.g., Dennis (2004) and Söderlind, Söderström, and Vredin (2005)), then we could still preserve the GMM framework by augmenting the system with appropriate first order conditions from the Lagrangian as shown by Söderlind (1999). Taking that step here, however, would amount to imposing false restrictions on the model given our assumption about the nature of policy. Generalizing the GMM algorithm to the cases of optimal commitment and discretion would certainly be desirable, but is properly the business of another paper.

Following Salemi (2006), we can rewrite the policymaker's loss function as

$$\Lambda = \mathrm{tr}\left(W \times (1-\delta)^{-1}\left[\Phi + \delta G \Phi G' + \delta^2 G^2 \Phi (G^2)' + \ldots\right]\right) = \mathrm{tr}\left(W \times M\right) \tag{5}$$

where M is the discounted sum of forecast error variances in X_{t+j} when policy is set at date t. The following closed form solution for M can be obtained by applying the *vec* operator:

$$\mathrm{vec}\left(M\right) = (1-\delta)^{-1}\left[I - \delta G \otimes G\right]^{-1} \mathrm{vec}\left(\Phi\right) \tag{6}$$

The optimal value of θ satisfies the first-order necessary conditions of the central bank's loss minimization problem given by

$$\frac{\partial \Lambda}{\partial \theta_k} = \mathrm{vec}\left(W\right)' \times \frac{\partial \mathrm{vec}(M)}{\partial \theta_k} = 0 \tag{7}$$

where θ_k corresponds to the k^{th} element of θ. Using (6) and recognizing that G depends on the policy-rule coefficients, we can obtain the following analytic expression for the partial derivative on the right-hand-side of (7):

$$\begin{aligned}
\frac{\partial \text{vec}(M)}{\partial \theta_k} &= \left(\frac{\delta}{1-\delta}\right)[I-\delta G \otimes G]^{-1}\left(\frac{\partial (G \otimes G)}{\partial \theta_k}\right)[I-\delta G \otimes G]^{-1} \times \text{vec}(\Phi) \qquad (8)\\
&= D_k(\rho, \theta) \times \text{vec}(\Phi)
\end{aligned}$$

where terms involving the Kronecker product of G are combined in the $(n_x^2 \times n_x^2)$ matrix D_k for notational convenience. In the construction of (8), we have assumed that Φ is unaffected by the choice of θ. One of the models discussed below, however, requires that we relax this assumption and modify the partial derivative expression accordingly.

2.1. Imposing optimality in the course of estimation

Given a sample $\{X_t\}_{t=1}^T$, the econometrician seeks estimates of the structural parameters subject to the auxiliary condition that the elements of θ are those that minimize the central bank's expected loss function (4). Salemi (2006) uses a "brute force" strategy to impose the auxiliary restriction. For given initial values of ρ and W, the brute force approach identifies the value of θ that minimizes expected loss, solves the model for its reduced form (3), and then computes sample log likelihood. The algorithm searches over values of ρ, W, and all unique elements of Φ that increase sample likelihood and stops when no higher value can be found. Estimation time can be long because the algorithm calculates the optimal policy each time log likelihood is recorded, that is, for many parameter values very different from those that fit the data.

The alternative approach developed here exploits the set of orthogonality conditions provided by the central bank's optimization problem to estimate the model using GMM. Specifically, the algorithm combines the least squares normal equations given by $E(\varphi_t X'_{t-1}) = 0$ with a collection of theoretical moments obtained by taking the unconditional expectation of (7). Denote $\hat{\varphi}_t$ the sample estimate of φ_t and let $\hat{\Phi} = 1/T \sum_{t=1}^T \hat{\varphi}_t \hat{\varphi}'_t$ be the sample average of the matrix of time t residual covariances. Using $\hat{\Phi}$ as an estimate of Φ, one can construct the sample analog of (7) as

$$\frac{\partial \Lambda}{\partial \theta_k} = \mathrm{vec}(W)' \times D_k(\rho, \theta) \times \mathrm{vec}(\hat{\Phi}) = 0 \tag{9}$$

for all k in θ. Equation (9) indicates that if the policy rule is indeed optimal, a certain linear combination of the elements of $\hat{\Phi}$ vanishes.

Define $g(\rho, \theta, W)$ to be the $(m \times 1)$ vector that contains the sample counterparts of the least-squares normal equations as well as the k first order conditions satisfied by an optimal policy rule. The typical element of g should be zero provided the model is true. The GMM estimation criterion is $Q = g(\rho, \theta, W)'S^{-1}g(\rho, \theta, W)$, where S^{-1} is the optimal weighting matrix described in Hamilton (1994, pp. 412-13). The algorithm searches freely over values of ρ, θ, and W for those that reduce Q and stops when no lower value can be found. Although the empirical strategy adopted here increases the estimated parameter space by the dimension of θ, it can reduce computation time relative to the brute force approach by avoiding the calculation of an optimal policy for each set of parameters evaluated.

3. GMM estimation of New Keynesian models

In what follows, we test the performance of the GMM algorithm formalized in section 2. By means of Monte Carlo simulations, we estimate the structural parameters of three different models subject to the condition that the policy equation minimizes a well-defined loss function. All three models are New Keynesian in spirit in that policy affects aggregate demand through a conventional interest rate channel and inflation through a Phillips curve specification. Each model determines the equilibrium relationship among the output gap y, the inflation rate π, and the short-maturity nominal interest rate r controlled by the central bank. While sharing some broad characteristics, the three models differ substantially in several important ways, namely, in the complexity of the relationship between structural and reduced-form parameters, in how the partial derivatives of expected loss are computed, and in the number of over-identifying restrictions they imply.

3.1. A backward-looking model

We begin by applying GMM to a purely backward-looking model. Although it can be criticized for providing no explicit role for expectations, it is instructive to start with a simple framework for a number of reasons. First, analytic expressions for the elements of $\frac{\partial(G \otimes G)}{\partial \theta_k}$ are easily obtained in the context of a backward-looking model. Making use of them to evaluate the sample moment conditions will likely render estimation more accurate. Second, the loss-minimizing values of θ can be verified independently by iterating on the matrix Ricatti equations. Third, the chosen backward-looking structure implies exact identification of the model parameters, whereas estimation of more elaborate forward-looking models entails over-identification.

The model consists of three equations that jointly govern the dynamics of the output gap, inflation, and the nominal interest rate. It is similar to the models used by Svensson (1997) and Ball (1999) to evaluate alternative targeting policies.

$$y_t = a y_{t-1} - b(r_t - \pi_t) + u_t \tag{10}$$

$$\pi_t = \alpha \pi_{t-1} + \beta y_t + v_t \tag{11}$$

$$r_t = \theta_y y_{t-1} + \theta_\pi \pi_{t-1} + w_t \tag{12}$$

Equation (10) is an IS schedule establishing output as a function of its own lag, a pseudo-real interest rate, $r_t - \pi_t$, and a white-noise demand shock u_t. All parameters are assumed positive. Equation (11) is a Phillips curve that illustrates the dependence of inflation on past inflation, a serially uncorrelated supply shock v_t, and a measure of excess demand which is assumed proportional to the output gap.

The central bank directs policy towards stabilizing a collection of target variables summarized by the following quadratic loss function:

$$\Lambda = E_t \sum_{j=0}^{\infty} \delta^j \left[\pi_{t+j}^2 + W_y y_{t+j}^2 + W_r r_{t+j}^2 \right] \tag{13}$$

A sequence $\{r_{t+j}\}_{j=o}^{\infty}$ corresponding to a unique combination of θ_y and θ_π is chosen so as to minimize (13) subject to (10) - (11). The three arguments contained in the loss function indicate that the monetary authority penalizes departures of inflation, the output gap, and the nominal interest rate from their respective target levels.[2] The relative size of the penalty attached to each is determined by the nonnegative policy weights $\{1, W_y, W_r\}$.

Loss functions of the form (13) are common in the monetary policy literature. Svensson (1999) argues that the principle stabilization objectives of an inflation targeting central bank can be represented by a weighted sum of the variances of target-adjusted inflation and the output gap. Under certain conditions, Woodford (2003) demonstrates that such a loss function can also be obtained as a quadratic Taylor series expansion of the expected utility of the representative consumer in an optimization-based model of the type examined later in this paper. Perhaps less common is the inclusion of an explicit interest rate stabilization objective. However, Woodford (2003) shows that a concern for nominal interest rate volatility can be justified on the grounds that it prevents frequent violations of the zero lower bound.

We should note that from an econometric standpoint, penalizing squared deviations of inflation and the nominal interest rate from constant trends may be problematic in practice considering the structural breaks that have likely occurred in the historical series from shifts in monetary policy regime. Dealing with structural breaks, however, is beyond the scope of this paper. As a result, we assume that the sample covers a period over which one can

[2] Because π and r are expressed as percent deviations from trend and y corresponds to the output gap, it is reasonable to assume that the appropriate target values for each of these variables is zero.

justify a stable policy regime.[3] In addition, a loss function penalizing first differences in the nominal interest rate (e.g., Rudebusch and Svensson (1999)) may capture the smoothing behavior of the central bank better than the present version. Because our main objective is to demonstrate the consistency of GMM, we leave such tasks for future work.

The parameter values chosen for the backward-looking model have a straightforward interpretation and are listed along with the estimation results in Table 1. To parameterize the reduced-form error covariance matrix Φ, we use estimates reported in Salemi (2006) which indicate that interest rate shocks are positively correlated with innovations to aggregate demand and supply. Concerning the loss function parameters, we assume that the central bank places a larger emphasis on stabilizing inflation than on stabilizing output while applying an intermediate weight on attaining interest rate stability. A policy rule of the form (12) with $\theta_y = 0.306$ and $\theta_\pi = 0.102$ minimizes expected loss.

In assessing the performance of the estimation algorithm, a key issue is determining whether or not GMM can successfully recover the optimal values of the policy-rule coefficients together with the loss function weights. Figure 1 provides some evidence on this matter, illustrating how the partial derivatives of Λ with respect to θ_y and θ_π vary with departures of θ_y from its optimal value.[4] For $\theta_y = 0.306$, the partial derivative functions return values on the order of 10^{-17}. As θ_y moves away from it loss-minimizing value, the partial derivatives increase rapidly to values ranging between 10^{-4} and 10^{-2}. Thus, an estimation criterion that includes first order conditions from the policymaker's control problem should be able to distinguish between optimal and suboptimal values of θ_y and θ_π.

Before proceeding further, we address an important econometric issue concerning the specific form of the moment conditions employed during estimation. It is well known that the properties of the ordinary least squares estimator can be obtained as a special case of

[3]See Salemi (2006) for a discussion on how to deal with structural breaks in the context of an inverse control exercise.

[4]The graph corresponding to variations in θ_π conveys similar information and is not displayed.

GMM. As a result, least-squares normal equations requiring zero covariance between the residuals and the regressors are frequently used as a basis for GMM. There are six such equations for the backward-looking model: $E(y_{t-1}\varphi_{j,t}) = 0$ and $E(\pi_{t-1}\varphi_{j,t}) = 0$ for $j = 1$, 2, and 3. In test estimations where the sample counterparts of the normal equations were used in conjunction with the policy optimality restrictions given by (9), GMM repeatedly converged to a particular set of perverse values: $a = \alpha = 1.0$ and $b = \beta = 0.0$. For these values, interest rate adjustments have no impact on the output gap or inflation, implying that any set of policy-rule coefficients will satisfy the first order conditions. When (9) was excluded from the estimation criterion, however, GMM consistently returned unbiased estimates that converged to the true values as sample size increased.

We suspect that the algorithm finds a local minimum at false parameter values when the normal equations are coupled with first order conditions. One potential explanation for this finding is that the numerical scale of the two sets of orthogonality conditions diverge. To illustrate this possibility, it is instructive to compute the population moments employed in GMM. Recall that G is the true reduced-form coefficient matrix and denote $\hat{G}$ an estimate of G. The least squares normal equations evaluated at $\hat{G}$ are

$$E\left[(X_t - \hat{G}X_{t-1})X'_{t-1}\right] = (G - \hat{G})\Phi_x = 0 \tag{14}$$

where Φ_x is the population covariance matrix of X_t. The same value of $G - \hat{G}$ implies a smaller value of Q, the GMM estimation criterion, if the diagonal elements of Φ_x fall.

As a remedy, we remove the dependence of the normal equations on the scale of the data by restating them as *correlations* rather than covariances. Using $\hat{\varphi}_t$ as an estimate of φ_t, one can construct the sample analog of the correlation restrictions as

$$\operatorname{corr}(y_{t-1}, \hat{\varphi}_{j,t}) = \frac{\sum_{t=1}^{T} y_{t-1}\hat{\varphi}_{j,t}}{\sqrt{\left(\sum_{t=1}^{T} y_{t-1}^2\right)\left(\sum_{t=1}^{T} \hat{\varphi}_{t,j}^2\right)}} = 0 \tag{15}$$

$$\operatorname{corr}(\pi_{t-1}, \hat{\varphi}_{j,t}) = \frac{\sum_{t=1}^{T} \pi_{t-1}\hat{\varphi}_{j,t}}{\sqrt{\left(\sum_{t=1}^{T} \pi_{t-1}^2\right)\left(\sum_{t=1}^{T} \hat{\varphi}_{t,j}^2\right)}} = 0 \tag{16}$$

for $j = 1$, 2, and 3. In test estimations where (15) - (16) were used in place of the conventional normal equations, GMM delivered unbiased estimates of the structural parameters that converged to the true values with sample size. As a result, the correlation versions of the least-squares normal equations were used to produce the estimates for all three models.

To evaluate the performance of GMM with the auxiliary moment restrictions, we conduct a battery of Monte Carlo experiments using two different parameterizations of the policy equation. The first set of parameters correspond to the optimal values of the policy-rule coefficients. The second set, however, is not optimal for any possible combination of loss function weights.[5] The data generating process for each is given by (3) where G is computed for the true parameter values and φ_t is the output of a multivariate normal random number generator.

For the backward-looking model, GMM estimation entails eight moment restrictions.[6] Six of those restrictions require that the residuals from each of the three reduced-form equations be uncorrelated with the lagged state variables $\{y_{t-1}, \pi_{t-1}\}$. The remaining two restrictions

[5]The suboptimal policy coefficients chosen for this particular exercise are $\theta_y = 0.20$ and $\theta_\pi = 2.00$. Given the structural parameters, which are held fixed across both cases, there are no values of W_y and W_r that render the policy rule optimal.

[6]The numerical routine used in the experiments was PATERN from Version 6 of the GQOPT Library of Fortran optimization programs. PATERN is a direct search algorithm that combines exploratory searches parallel to the parameter-space axes. As the performance of direct search algorithms is known to be sensitive to initial step size, PATERN was called several times in succession with decreasing initial step sizes. The estimation algorithm employed two sets of calls to PATERN. For the first set, the GMM weighting matrix was the identity matrix. The optimal weighting matrix was then estimated according to the formula given by Hamilton (1994, p. 413) before a second set of calls to PATERN was undertaken.

require that the partial derivatives of Λ with respect to θ_y and θ_π vanish.

Table 1 reports estimates for the case where the policy-rule coefficients are loss minimizing. The left panel reports results when the optimal-policy restrictions are not imposed during estimation. The right panel reports results when the restrictions are imposed. The left panel reports statistics for six parameters $\{a, b, \alpha, \beta, \theta_y, \theta_\pi\}$; the right panel for these six plus $\{W_y, W_r\}$. The parameters are exactly identified in both panels.

The typical entries in the table are the average and standard deviation of parameter estimates computed across a subset of 100 trials. Trials where estimates converged to outlying values were excluded from the statistics on the grounds that a researcher would re-start the algorithm rather than accept the outlying estimates. The table reports the fraction of trials over which average and standard deviation were computed. As sample size increased, fraction converged to 1.0.

Table 1 supports a number of conclusions. First, GMM returns unbiased estimates of all structural parameters that converge to the true values as sample size increases. The policy-rule coefficients are also unbiased and precisely estimated even in small samples. A comparison across both panels illustrates that the consistency of the GMM estimator is unchanged by the inclusion of auxiliary moment restrictions that constrain the choice of θ. Second, when the optimality restrictions are imposed, GMM delivers unbiased estimates of W_y and W_r that converge to the true values with sample size. Unlike the remaining structural parameters, however, convergence of the loss function weights is slower, resulting in estimates that are statistically insignificant in smaller samples. Although not reported in the table, the partial derivatives of loss with respect to the policy coefficients averaged 10^{-6} in samples of size 250 and 10^{-10} in samples of size 5000. We conclude that augmenting the GMM criterion with the appropriate optimality restrictions is a useful way of estimating policy-rule coefficients that minimize expected loss.

Table 2 reports the case when the optimal policy restrictions are imposed even though the true values of the policy-rule coefficients are not loss minimizing for any values of $\{W_y, W_r\}$. Imposing false optimality restrictions does not bias estimates of a, b, or α, and does not bias estimates of the policy-rule coefficients. It seriously biases estimates of β, the coefficient on the output gap in the Phillips curve. It also produces estimates of W_y and W_r that converge to zero with sample size. As we will see, the finding that including false optimality conditions biases estimates of structural parameters but not policy-rule coefficients holds true in all of the models we study.

This exercise illustrates that caution should be taken when conditioning estimation on the assumption of policy optimality. For our backward-looking model, a casual interpretation of the results would lead to the erroneous conclusion that output-gap fluctuations have little impact on inflation. Moreover, estimates of W_y and W_r would imply that the central bank cares only about stabilizing inflation.

Why does our algorithm lead to bias of some structural parameters but not the policy-rule coefficients? We conjecture that the normal equations tightly identify the policy coefficients. Consequently, when forced to locate an optimal policy, GMM seeks values of the structural parameters and loss function weights that make the true value of θ appear optimal. Instead of settling on biased values of the policy coefficients, the algorithm finds an alternative economic world in which the true values of the policy coefficients are closer to optimal.[7]

3.2. A forward-looking rational expectations model

In this section we apply the GMM algorithm to the small-scale empirical New Keynesian model estimated by Salemi (2006). It is structurally similar to the kinds of models popularized by Clarida *et al.* (1999) in that the key aggregate relationships are compatible with an underlying framework based on optimizing agents. While it emphasizes the role of forward-

[7]We address this issue in more detail in an appendix to the paper which can be accessed along with other supplementary material at ScienceDirect (www.sciencedirect.com) or from the corresponding author upon request.

looking behavior and rational expectations, the model also incorporates a substantial degree of endogenous persistence in the form of multiple lags of output and inflation. The complete model is a three equation system given by

$$\begin{aligned} y_t &= \lambda E_t y_{t+1} + a_1 y_{t-1} + a_2 y_{t-2} - b(r_t - E_t \pi_{t+1}) + u_t && (17) \\ \pi_t &= \alpha_1 E_t \pi_{t+1} + \alpha_2 \pi_{t-1} + \beta y_t + v_t && (18) \\ r_t &= \theta_{y1} y_{t-1} + \theta_\pi \pi_{t-1} + \theta_r r_{t-1} + \theta_{y2} y_{t-2} + w_t && (19) \end{aligned}$$

where all variables carry the same definition used in the previous example and each is expressed as percent deviations from trend. The stochastic variables $\{u_t, v_t, w_t\}$ are serially uncorrelated shocks to aggregate demand, aggregate supply, and monetary policy, respectively.

The IS equation (17) is loosely consistent with a linearized Euler condition characterizing the optimal consumption plan in a dynamic general equilibrium setting. As explained in Clarida *et al.* (1999), the inverse relationship between current output and the real interest rate reflects intertemporal substitution on the part of households, and the presence of expected future output is motivated by a desire to smooth consumption.[8] In contrast, the rationale for including two lags of output is largely empirical. Fuhrer and Rudebusch (2004), for instance, obtain formal estimates of the parameters of a generalized New Keynesian output equation and conclude that multiple lags are essential for explaining the dynamic properties of real output.

Equation (18) is a "hybrid" version of the New Keynesian Phillips curve analyzed by Galí and Gertler (1999). The dependence of current inflation on expected future inflation and the output gap emerges from an environment of monopolistically competitive firms that adjust

[8]We say that (17) is "loosely" consistent with the consumption Euler condition because, as shown in the next section, an aggregate demand specification derived explicitly from household optimization implies a number of additional cross-parameter restrictions on the values of λ, a_1, a_2, and b.

prices in a staggered fashion (e.g., Taylor (1980) and Calvo (1983)). The presence of lagged inflation is justified largely on the basis of recent empirical studies. Estrella and Fuhrer (2002), for example, criticize the purely forward-looking Phillips curve on the grounds that it is inconsistent with the inertial behavior of inflation observed in U.S. data.

In the backward-looking model, one can easily find analytic expressions for the reduced-form matrix G. The addition of forward-looking variables and rational expectations in the present model, however, make the construction of analytic solutions problematic. We use the technique of Blanchard and Kahn (1980) to find the reduced-form solution to the system given by (17) - (19). Accordingly, we define the state vector to be $X_t = [y_t \ \pi_t \ r_t \ y_{t-1}]'$ and express the model in compact form as:

$$\begin{bmatrix} X_t \\ E_t y_{t+1} \\ E_t \pi_{t+1} \end{bmatrix} = B \begin{bmatrix} X_{t-1} \\ y_t \\ \pi_t \end{bmatrix} + D \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} \tag{20}$$

where B and D are (6×6) and (6×3) matrices the elements of which are completely determined by the set of underlying structural parameters and the policy rule coefficients. A unique bounded solution of the form (3) exists if the number of unstable eigenvalues of B equals the number of forward-looking variables in (20).[9]

Without analytic solutions, estimation becomes more complicated because there are no analytic expressions for the elements of $\frac{\partial(G \otimes G)}{\partial \theta_k}$. Consequently, we employ symmetric finite differences to obtain a numerical approximation of the partial derivative expression involving the Kronecker product of G. Figure 2 illustrates how the partial derivatives of Λ with respect to all four policy-rule coefficients vary with departures of θ_π from its optimal value. At the optimum, the numerical derivative function returns numbers on the order of 10^{-12}. The

[9]The values chosen for the structural parameters are taken from Salemi (2006) and satisfy all the necessary stability conditions. Those values along with the estimation results are reported in Table 3.

derivatives increase rapidly to values in the neighborhood of 10^{-3} as θ_π moves away from its optimal value. The implication is that an estimation criterion that includes (9) can still discriminate between optimal and suboptimal values of θ even when analytic expressions for the elements of G are not available.

To assess the performance of GMM in the context of a forward-looking rational expectations model, we conduct a battery of Monte Carlo experiments. When the optimal policy restrictions are not imposed, estimation is based on twelve least-squares normal equations. These include the sample correlations between the lagged state variables $\{y_{t-1}, \pi_{t-1}, r_{t-1}, y_{t-2}\}$ and the three reduced-form errors. When the optimality hypothesis is imposed, the GMM criterion includes the normal equations in addition to four partial derivative restrictions corresponding to $\{\theta_{y1}, \theta_\pi, \theta_r, \theta_{y2}\}$.

Table 3 reports findings for the case in which the optimality hypothesis is true. A number of conclusions can be drawn. First, GMM consistently returns unbiased estimates of all structural parameters that converge to the true values with sample size. Estimates of the policy rule are unbiased and statistically significant even in small samples. A comparison across both panels reveals that these results are unaffected by the use of optimality restrictions in the course of estimation. Second, although there is little evidence of bias, GMM tends to deliver imprecise estimates of some structural parameters in smaller samples. The standard error for α_2, for instance, is quite large for a sample size of 100. Third, imposing the optimality restrictions when they are true reduces the uncertainty surrounding the estimates of many key structural parameters. The improvement is most noticeable for the IS equation, as the sample standard errors accompanying the estimates of λ, a_1, a_2, and b are each smaller than their counterparts under least squares estimation. Fourth, when the optimality hypothesis is imposed, GMM returns unbiased estimates of W_y and W_r that converge to the true values with sample size.

Table 4 reports findings for the case in which a false optimality assumption is imposed.

In this example, the true values of the policy-rule coefficients ($\theta_{y1} = .50$, $\theta_\pi = 1.50$, $\theta_r = .50$, $\theta_{y2} = 0$) do not minimize the central bank's loss function for any combination of weights $\{W_y, W_r\}$. Despite conditioning estimation on false optimality restrictions, GMM still returns unbiased estimates of the policy-rule coefficients at all sample sizes. In contrast, the estimates of some structural parameters are biased and insignificant. The estimate of λ is far below the true value, implying a weak connection between current output and expected future real interest rates. The estimates of α_2 and β are likewise too small. These results would lead a researcher to the mistaken conclusion that the inflation process is less inertial and that fluctuations in excess demand have a more modest impact on inflation dynamics. Finally, estimates of W_y and W_r are near zero and statistically insignificant. Thus, given the observed policy behavior embodied by the actual feedback coefficients, basing estimation on a set of false optimality conditions drives the loss function weights to values that would suggest a policy of strict inflation targeting.

In contrast to the backward-looking model, estimation of our forward-looking model entails over identification of the structural parameters. When the optimality restrictions are not imposed, twelve moment conditions (all three residuals must be uncorrelated with each of the four regressors) are used to obtain estimates of eleven parameters. Imposing the optimality restrictions expands the parameter space by two (W_y and W_r) while adding four moment conditions (four partial derivatives). Because the number of orthogonality conditions exceeds the number of parameters to be estimated, we can test the restrictions implied by the forward-looking model under both assumptions about policy. Under the null hypothesis that the actual population moments are truly zero, Hansen (1982) proposes a simple test based on the finding that $Q \times T$ should be asymptotically distributed chi-squared with degrees of freedom equal to the number of over-identifying restrictions.[10]

Table 5 reports the rejection frequency of the over-identifying restrictions as a function

[10]Recall that Q is the minimized GMM criterion and T is the sample size.

of test size and sample size. When the optimality hypothesis is not imposed, estimation implies one over-identifying restriction. When it is imposed, estimation implies three over-identifying restrictions. Overall, the figures in Table 5 support three general conclusions. One, the likelihood of rejecting the over-identifying restriction is higher than expected in small samples when the optimality hypothesis is not imposed. The rejection frequency does, however, gradually converge to the expected number as sample size increases. Two, the chi-squared test rejects too often even in large samples when estimation is conditioned on the true hypothesis of policy optimality. At a sample size of 5000, for instance, the restrictions are rejected in forty-one percent of samples by a twenty-five percent test and in fourteen percent of samples by a one percent test. Thus, it appears that GMM will too often reject the over-identifying restrictions implied by the forward-looking model even when the corresponding population moments are really zero. Three, imposing a false optimality hypothesis leads to a rejection of the over-identifying restrictions at every test size and every sample size over 100. The implication is that the standard test has substantial power to reject the optimality restrictions when they are indeed false.

3.3. A representative-agent general equilibrium model

The third model selected for estimation belongs to a larger family of dynamic general equilibrium models described by Goodfriend and King (1997) as the "New Neoclassical Synthesis." The model integrates Keynesian elements, like staggered price-setting and monopolistic competition, into an otherwise standard business cycle framework emphasizing intertemporal optimization and rational expectations.[11] In short, a representative household chooses optimal sequences of consumption and labor supply to maximize expected lifetime utility subject to a conventional budget constraint. Profit-maximizing firms stagger price contracts in the fashion of Calvo (1983) and manufacture differentiated products using la-

[11] Models belonging to this family include Rotemberg and Woodford (1997), McCallum and Nelson (1999a), and King and Wolman (1999).

bor and capital. In addition, the model features two sources of mechanical persistence in the form of habit formation in consumption (e.g. Furhrer (2000)) and partial indexation to lagged inflation (e.g. Smets and Wouters (2003)).

The complete model characterizes the equilibrium dynamics of four variables: y_t, π_t, r_t, and y_t^n, the natural rate of output prevailing under flexible prices. The following equations constitute a log-linear approximation of the model's equilibrium conditions expanded around a zero-inflation steady state.

$$
\begin{aligned}
b\Delta y_t &= (1+\beta b^2)E_t\Delta y_{t+1} - \beta b E_t \Delta y_{t+2} - \tilde{\sigma}\left[r_t - E_t\pi_{t+1}\right] + \sigma^{-1}(1-b)u_t && (21)\\
\pi_t &= \frac{\gamma}{1+\beta\gamma}\pi_{t-1} + \frac{\beta}{1+\beta\gamma}E_t\pi_{t+1} + \left(\frac{(1-\varepsilon)(1-\beta\varepsilon)}{(1+\beta\gamma)\varepsilon}\right)\left(\frac{\chi+\alpha}{1-\alpha}\right)(y_t - y_t^n) && (22)\\
y_t^n &= \frac{1-\alpha}{\chi+\alpha}\left[\frac{1+\chi}{1-\alpha}v_t + (1-\beta b)^{-1}u_t - \tilde{\sigma}^{-1}\left[(1+\beta b^2)y_t^n - b y_{t-1}^n - \beta b E_t y_{t+1}^n\right]\right] && (23)\\
r_t &= \theta_\pi \pi_{t-1} + \theta_y y_{t-1} + \theta_r r_{t-1} + w_t && (24)
\end{aligned}
$$

where $\tilde{\sigma} \equiv \sigma^{-1}(1-b)(1-\beta b)$ and Δ is the first difference operator.[12]

Equation (21) can be interpreted as an intertemporal IS schedule where $\tilde{\sigma}$ measures the sensitivity of consumption plans to changes in the real interest rate. The stochastic parameter u_t is a serially uncorrelated demand shock generating exogenous variation in the marginal utility of consumption. As illustrated by Amato and Laubach (2004), habit formation implies that the current *growth rate* of output depends on expectations of future growth rates. Without habit formation ($b = 0$), (21) collapses to the familiar IS equation discussed in Woodford (1999).

Equation (22) is a Phillips curve governing the dynamic behavior of inflation. The assumption that firms index to lagged inflation when they are blocked by the Calvo mechanism from re-optimizing their price makes current inflation depend on past inflation. Without in-

[12]Details about the preference structure of the model and the corresponding equilibrium conditions can be found in an appendix to the paper available from ScienceDirect.

dexation ($\gamma = 0$), (22) reduces to the purely forward-looking New Keynesian Phillips curve analyzed by Galí and Gertler (1999) that links inflation to expected future inflation and the theoretical output gap defined as $y_t - y_t^n$.

Because the loss function consistent with the present model depends on $y_t - y_t^n$, it is necessary to track the evolution of output under flexible prices. Equation (23) implicitly defines y_t^n as a function of y_{t-1}^n and two stochastic disturbances, the demand shock u_t and a serially uncorrelated technology shock v_t.

In the spirit of Rotemberg and Woodford (1997), alternative policies are ranked on a welfare-basis according to a loss function that is derived by taking a quadratic approximation to the representative consumer's expected lifetime utility.

$$\Lambda = E_t \sum_{j=0}^{\infty} \beta^j \left[W_\pi (\pi_{t+j} - \gamma \pi_{t+j-1})^2 + W_y \left((y_{t+j} - y_{t+j}^n) - \delta_y (y_{t+j-1} - y_{t+j-1}^n)^2 \right) \right] \qquad (25)$$

The added persistence generated by indexation and habit formation implies that the policy goals consistent with household optimization involve stabilizing a measure of inflation relative to its own lag and the current output gap relative to last period's. Additionally, the set of coefficients $\{W_\pi, W_y, \delta_y\}$ are not free, but rather specific functions of the underlying structural parameters. In the absence of indexation ($\gamma = 0$) and habit formation ($b = 0 \Rightarrow \delta_y = 0$), (25) reduces to the well-known loss function defined over the second moments of inflation and the output gap alone.[13]

To find the rational expectations solution to the system of equations given by (21) - (24), we re-define the state vector to be $X_t = [y_t \ \pi_t \ r_t \ y_t^n \ y_{t-1} \ \pi_{t-1} \ y_{t-1}^n]'$ and express the model

[13]For a comprehensive derivation of the welfare function, refer to Woodford (2003, Chapter 6)

in compact form.

$$\begin{bmatrix} X_t \\ E_t y_{t+1} \\ E_t y_{t+2} \\ E_t \pi_{t+1} \end{bmatrix} = B \begin{bmatrix} X_{t-1} \\ y_t \\ E_t y_{t+1} \\ \pi_t \end{bmatrix} + D \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix} \tag{26}$$

B and D are (10×10) and (10×3) matrices whose elements are nonlinear functions of the parameters appearing in the model. Having augmented the state vector with the relevant lags, we can reformulate (25) in terms of X_t with weight matrix given by

$$W = \begin{bmatrix} W_y & 0 & 0 & -W_y & -\delta_y W_y & 0 & \delta_y W_y \\ 0 & W_\pi & 0 & 0 & 0 & -\gamma W_\pi & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -W_y & 0 & 0 & W_y & \delta_y W_y & 0 & -\delta_y W_y \\ -\delta_y W_y & 0 & 0 & \delta_y W_y & \delta_y^2 W_y & 0 & -\delta_y^2 W_y \\ 0 & -\gamma W_\pi & 0 & 0 & 0 & \gamma^2 W_\pi & 0 \\ \delta_y W_y & 0 & 0 & -\delta_y W_y & -\delta_y^2 W_y & 0 & \delta_y^2 W_y \end{bmatrix}.$$

We use the method of Blanchard and Kahn (1980) discussed in the previous section to determine the model's reduced-form representation.

$$X_t = G X_{t-1} + H \varepsilon_t \tag{27}$$

G and H are (7×7) and (7×3) matrices of reduced-form coefficients, and $\varepsilon_t = [u_t \; v_t \; w_t]'$ is the vector of structural shocks with covariance matrix Σ.[14]

Estimating a model that is specified at the level of individual preferences presents challenges that do not emerge in the previous two models. For one, the mapping from the

[14]In terms of the notation introduced earlier, $H\varepsilon_t = \varphi_t$, implying that structural and reduced-form error covariance matrices are related by $\Phi = H\Sigma H'$.

structural parameters to the reduced form is more complicated because the slope coefficients appearing in (21) - (23) are themselves nonlinear functions of the structural parameters. This makes it impossible to identify every parameter, and as a result, some must be fixed prior to estimation. For parameters that are identified, the complexity of the additional cross-equation restrictions sometimes makes it difficult to obtain precise estimates in small samples. The identified parameters include the degree of habit formation b, the inverse elasticity of substitution σ, the fraction of non-adjusting firms ε, the inverse wage elasticity of labor supply χ, and the degree of partial indexation γ. A description of each parameter, including those that are not identified, appears in Table 6. The values chosen for the structural parameters are broadly consistent with estimates obtained in numerous empirical studies of dynamic general equilibrium models.

A second challenge arises due to the relationship between the structural parameters and the reduced form errors. Recall that the exogenous shocks have a particular economic interpretation within the context of a representative agent model. Consequently, the *structural* error covariance matrix Σ should remain invariant to changes in the structural parameters ρ and the policy rule coefficients θ. If Σ is fixed, however, a change in ρ or θ implies a change in the *reduced-form* error covariance matrix given by Φ. It follows that the partial derivatives of Λ with respect to the elements of θ must account for the implied change in Φ in order to correctly compute the first order conditions associated with an optimal policy. Recognizing that changes in θ now impact G and Φ, we obtain the following modification of the partial derivative expression appearing in (7):

$$\frac{\partial \mathrm{vec}(M)}{\partial \theta_k} = D_k(\rho, \theta) \times \mathrm{vec}(H\Sigma H') + \left(\frac{1}{1-\beta}\right)[I - \beta G \otimes G]^{-1} \times \frac{\partial \mathrm{vec}(H\Sigma H')}{\partial \theta_k} \qquad (28)$$

where D_k is the matrix defined in (8) with the central bank discount factor given by β. Denote $\hat{\varepsilon}_t$ the sample estimate of ε_t which can be recovered from the estimate of $\hat{\varphi}_t$. Let

$\hat{\Sigma} = 1/T \sum_{t=1}^{T} \hat{\varepsilon}_t \hat{\varepsilon}_t'$ be the corresponding matrix of time t residual variances. Using $\hat{\Sigma}$ as an estimate of Σ in the modified partial derivative expression, one can construct the sample analog of the central bank's first order conditions summarized by (7).

We assess the performance of the GMM algorithm by Monte Carlo simulations. When the optimality hypothesis is not imposed, the estimation criterion is based on nine least-squares normal equations. Specifically, these include the sample correlations between $\{y_{t-1}, \pi_{t-1}, r_{t-1}\}$ and the three reduced-form errors associated with output, inflation, and the nominal interest rate. When the optimality hypothesis is imposed, estimation is based on twelve restrictions, the nine normal equations plus three partial derivative restrictions corresponding to $\{\theta_\pi, \theta_y, \theta_r\}$. In contrast to the previous two examples, the assumption that the policymaker minimizes a loss function consistent with household welfare implies that the preference weights are known functions of the structural parameters. Thus, imposing the optimality hypothesis increases the number of moment conditions used for estimation while leaving the number of parameters to be estimated unchanged.

Table 7 reports findings for the case in which the optimality hypothesis is true. The experiments support several findings. First, for all structural parameters that can be identified, GMM returns unbiased estimates that converge to the true values with sample size. This result does not hinge on whether or not the optimality restrictions are imposed in the course of estimation. Second, GMM delivers imprecise estimates of σ and χ when the sample size is relatively small. Inspection of the model reveals that σ is identified through the impact of changes in the real interest rate on output growth in (21), and χ through the affect of fluctuations in the theory-based output gap on inflation in (22). Unfortunately, both parameters are confounded with others that are estimated with more precision, making them difficult to identify in small samples. Third, imposing the optimality restrictions when they are true sharpens estimates of many of the structural parameters. The standard errors for γ, ε, and χ are an order of magnitude smaller than their counterparts under least-squares estimation in large samples. Fourth, the ability of GMM to deliver unbiased estimates of the structural parameters guarantees that the weights in the central bank's objective function converge to the true values with sample size.[15]

Concerning the policy-rule coefficients, GMM returns unbiased estimates of $\{\theta_\pi, \theta_y, \theta_r\}$ at all sample sizes regardless of whether or not the optimality restrictions are imposed. In contrast to many of the structural parameters, the policy coefficients are precisely estimated even for small samples. Evidently, the primary advantage in the present model of basing estimation on an expanded set of moment conditions is that it reduces the uncertainty surrounding some of the key structural parameters.

Table 8 reports findings for the case in which the optimality hypothesis is false. The Monte Carlo evidence suggests that GMM consistently returns unbiased estimates of the policy-rule coefficients at all sample sizes. Imposing false optimality restrictions, however, leads to biased estimates of several important structural parameters. Like with the previous models, the normal equations place a stronger set of restrictions on the policy coefficients than they do on the structural parameters. Thus, when estimation is conditioned on a false optimality assumption, GMM basically searches for values of the structural parameters that render the true policy coefficients optimal. In other words, the algorithm tries to locate an alternate economic universe in which the observed policy rule would be nearly optimal. The outcome is biased estimates of the structural parameters but unbiased estimates of the policy rule-coefficients.

Table 9 presents the frequency of rejection of the over-identifying restrictions as a function of test size and sample size. When the optimality hypothesis is not imposed, estimation implies one over-identifying restriction. When it is imposed, estimation implies four over-identifying restrictions. The figures in Table 9 indicate that the rejection frequency is too high

[15] The values of W_π and W_y reported in the Tables 7 and 8 are the ones implied by the estimates of the remaining structural parameters. The sample standard errors are computed in the usual way.

in small samples when estimation is based on the least-squares normal equations alone. As sample size increases, however, the rejection frequency converges to the expected number. Similarly, the likelihood of rejection is too large in smaller samples when the optimality restrictions are true and imposed. Finally, in the event that a false optimality hypothesis is imposed, the over-identifying restrictions are rejected at every test size and every sample size over 100. Even at a sample size of 100, the restrictions are rejected in 88 percent of samples by a one percent test and in 98 percent of samples by a five percent test. Thus, the test again demonstrates great power to reject the optimal-policy moment restrictions when they are false.

4. Taking the models to the data with GMM

In this section we demonstrate our GMM algorithm with an application to U.S. data. The sample includes quarterly observations spanning 1979:III to 2001:IV on the output gap, the target-adjusted inflation rate, and the target-adjusted nominal interest rate.[16] We use the first two observations to initialize the system and treat the remaining data as observations on a single policy regime. To underscore the significance of imposing optimal policy restrictions in the course of estimation, we compare outcomes in two cases. In the first case, reaction function coefficients are unrestricted, and in the second, they are restricted to satisfy the first order conditions for a loss-minimizing policy.

We begin with an assessment of model fit. For all three models and for both assumptions about policy, Table 10 reports the minimized GMM criterion, the chi-squared test statistic for the model's over-identifying restrictions, and "pseudo log likelihood" obtained by scaling the natural logarithm of the determinant of the residual-error covariance matrix.

Table 10 supports a number of conclusions. First, the low pseudo log likelihood value indicates that the backward-looking model fits the data poorly. This finding is not sur-

[16] Refer to Salemi (2006) for information about the data set and for a discussion of the detrending procedure.

prising considering the abbreviated lag length and the absence of any transmission lags in the model. Rudebusch and Svensson (1999) demonstrate that a purely backward-looking model can fit the data with multiple lags in the IS and Phillips curve equations. Second, the forward-looking model fits the data better than the representative agent model with or without optimal policy restrictions imposed. Nevertheless, one would reject the null hypothesis that observed policy-rule coefficients are those that minimize expected loss at standard significance levels. Third, the representative agent model describes the data well when optimal policy restrictions are not imposed. The hypothesis that the Federal Reserve maximizes expected utility, however, is soundly rejected.

Table 11 reports estimates of the structural parameters for the forward-looking and representative agent models.[17] The estimates for the forward-looking model are properly signed, of plausible size, and close to the values reported in Salemi (2006). Imposing optimal policy restrictions alters some of the parameter estimates in a meaningful way. For example, the estimate of λ becomes smaller, pointing to a reduced role for expected output in the IS equation when policy-rule coefficients are forced to satisfy the necessary conditions for loss minimization. At the same time, the estimates of a_1 and a_2 become larger in magnitude, signalling an increased dependence on lagged output. Interestingly, the estimates of the Phillips curve are little affected by the inclusion of optimal policy restrictions. In both cases, the estimate of α_1 suggests that inflation is primarily backward looking, and the estimate of β indicates that the Phillips curve is relatively flat.

Imposing optimal policy restrictions has an even larger impact on estimates of the representative agent model. When the reaction function coefficients are unrestricted, the estimate of b, the degree of habit formation, is 0.87, close to the values reported by Fuhrer (2000)

[17] The parameter estimates for the backward-looking model have the wrong sign in many cases. Rather than re-specifying the model in a way that improves its empirical fit, we concentrate on estimates of the forward-looking and representative agent models. Both models are more richly parameterized and have been taken to the data by others.

and Heaton (1995). In the restricted case, however, the estimate of b approaches the upper limit of the parameter space. The coefficient of relative risk aversion σ is 4.43 when the optimality restrictions are not imposed and near zero when they are. Regarding the estimates that govern price-setting behavior, the indexation parameter γ is 0.76 and the Calvo probability ε is 0.98 when policy-rule coefficients are unrestricted. Both of these estimates are somewhat larger, but still consistent with, the values reported by Smets and Wouters (2005). Augmenting GMM with optimal policy restrictions, however, lowers the degree of price stickiness and drives the indexation parameter to unity. Finally, the inverse labor supply elasticity parameter χ is near zero in both cases, implying that the labor supply schedule is essentially flat.

Table 11 also reports estimated standard errors for both models computed using the formula for the asymptotic covariance matrix given in Hamilton (1994, p. 415). The estimated standard errors are large in many cases, indicating that the data contain only imprecise information about the structural parameters. However, standard errors are typically smaller when the optimal policy restrictions are imposed, bolstering our claim that incorporating moment conditions consistent with loss minimization reduces uncertainty surrounding many parameter estimates. The exceptions are those coefficients that reside on the boundary of the allowable parameter space.

Table 12 reports estimates of the policy-rule coefficients and the corresponding loss function weights for the forward-looking and representative agent models. To help clarify the impact of using optimality conditions in the course of estimation, we also report the partial derivatives of loss for both the unrestricted and restricted cases.[18] For the forward-looking model, imposing optimal policy restrictions leads to larger estimates of the coefficients on lagged output, illustrating the tension between fitting the interest rate series and satisfy-

[18]For the forward-looking model, we use the weights estimated in the restricted case to compute derivatives. For the representative agent model, we use the weights implied by the unrestricted estimates of the structural parameters.

ing the loss-minimizing criteria. The coefficient on the lagged interest rate is large and significant in both cases, indicating that policy inertia is a characteristic shared by actual Federal Reserve behavior as well as policies that minimize expected loss. The coefficient on lagged inflation is also positive and significant regardless of the inclusion of optimal policy restrictions. Taken together, the results imply that a one percent increase in inflation is accompanied by an even greater percent increase in the interest rate over time as required by the Taylor principle. Finally, the data are best reconciled with the optimality hypothesis under a loss function that places all weight on inflation stabilization and none on output gap or interest rate stabilization.[19]

The policy-rule coefficients for the representative agent model are similar to those of the forward-looking model when the optimality conditions are not imposed during estimation. The coefficients on lagged inflation and the lagged interest rate, for instance, are positive and significant. The coefficient estimate on lagged output is approximately equal to the sum of the coefficients on one and two lags of output in the forward-looking model. Based on the derivative estimates, it is also clear that the reaction function coefficients are far from those that maximize expected utility. The partial derivatives of loss are on the order of 10^1 to 10^3. Given the estimates of the structural parameters in Table 11, the quadratic welfare function discussed in section 3.3 places a much greater weight on quasi-differenced inflation than on the quasi-differenced output gap.

The results are dramatically altered when optimal policy restrictions are imposed during estimation. The coefficient on lagged output is a much larger negative, the inflation coefficient is very large and positive, and the coefficient on the lagged interest rate is close to zero. None of these is estimated with much precision. While smaller than their unrestricted counterparts, the partial derivatives are not close enough to zero to make certain that GMM

[19]Salemi (2006) reports a similar finding. Favero and Rovelli (2003) and Dennis (2004) also report very little concern for output gap stability in U.S. data.

has located an optimal policy. Thus, we conclude that the estimation algorithm cannot reconcile U.S. data with the hypothesis that the Federal Reserve set policy to maximize expected utility. In fact, we find that the parameter vector that fits the data best violates the saddlepath property, the eigenvalue condition that ensures a stable, unique rational expectations equilibrium. We take this as additional evidence that imposing optimal policy restrictions leads to a severe deterioration in the empirical performance of the representative agent model.

5. A comparison of GMM with alternative procedures

To better illustrate the potential costs and benefits of the GMM algorithm, we compare outcomes using our methodology with outcomes from two alternative estimation procedures commonly used in the inverse control literature. The first approach is the nested or "brute force" strategy discussed in the introduction which applies full information maximum likelihood (FIML) to estimate the structural parameters of the model as well as the loss function weights. The second approach employs a two-step estimator based on GMM methods that has recently been examined by Lippi and Neri (2007) and Ehrmann and Smets (2003) in the context of a small-scale New Keynesian model. In discussing the differences among the three estimation procedures, we perform Monte Carlo experiments only on the forward-looking model studied in section 3.2.

5.1. GMM vs. FIML

In contrast to our GMM strategy, the FIML procedure nests the loss minimization problem of the central bank within the estimation algorithm of the econometrician. Formally, the nested approach searches over values of the structural parameters and loss function weights for those that maximize the likelihood function implied by the model's reduced form (e.g., Dennis (2004), Dennis (2005), and Salemi (2006)). The optimal coefficients of the policy rule are computed for each set of parameters considered in the course of estimation, including those that do not fit the data.[20]

In comparing the performance of the estimation procedures, we conduct two separate Monte Carlo experiments. In the first experiment, the true policy-rule coefficients are optimal and, in the second, they are not optimal for any parameterization of the loss function. Table 13 reports findings for the case in which the optimality hypothesis is true. Not surprisingly, we find that FIML and GMM deliver unbiased estimates of all model parameters that approach the true values as sample size increases. Although we argue that GMM is computationally more efficient, FIML appears to have at least one advantage.[21] In small samples the FIML standard errors are uniformly smaller than their GMM counterparts except for those associated with the policy-rule coefficients (in which case they are approximately equal).

Table 14 reports findings for the case in which the optimality hypothesis is false. Like GMM, imposing false optimality restrictions leads to bias in the FIML estimates of some structural parameters. However, the bias need not appear in the same coefficients for both procedures. For example, the GMM estimate of λ is biased towards zero but the FIML estimate is unbiased. Alternatively, the FIML estimate of α_1 is too large while the GMM estimate is accurate. Our methodology does have one clear advantage over the nested approach. When false optimality restrictions are imposed, the FIML estimates of the policy-rule coefficients are significantly biased and the standard errors are large. In contrast, the GMM estimates are always unbiased and precise irrespective of the assumptions made about central bank preferences.

5.2. GMM vs. a two step estimator

The finding that policy-rule coefficients are always estimated without bias and with good

[20]Refer to the appendix and other supplementary files (available from ScienceDirect) for a more detailed discussion of the FIML procedure as well as a comparison of the estimation results with GMM.

[21]The appendix contains a detailed discussion about the computation time required by both procedures when the optimality hypothesis is true and when it is false.

precision under GMM raises the possibility that a two step estimation procedure could generate more favorable outcomes. We address this point by estimating only the structural parameters and the policy-rule coefficients in a first step using moment conditions based on the least-squares normal equations alone. In a second step, we estimate the loss function weights using a GMM criterion formed exclusively by the partial derivatives of expected loss while holding all other parameters fixed at their first stage values.[22]

Table 15 displays the results of Monte Carlo experiments of the forward-looking model under the proposed two step approach as well as the unified procedure described in section 2. It is clear that the two step procedure yields unbiased estimates of all parameters that converge to the true values with sample size, including the loss function weights obtained in the second stage estimation. The standard errors, however, tend to be larger than those estimated under the unified approach. The loss of precision occurs because the partial derivatives of expected loss convey information that helps identify structural parameters when the optimality hypothesis is true.

Despite the accuracy of the parameter estimates, Monte Carlo evidence reveals that the two step procedure can deliver misleading results about the true conduct of policy. Table 16 reports the partial derivative estimates of expected loss with respect to the policy-rule coefficients for the unified and two step procedures. For sample sizes of 100 and 1,000, the partial derivative estimates are very large. Only at a sample size of 10,000 do the two-step partial derivatives approach zero. As a result, standard tests of the loss-minimizing restrictions based on the two step estimates will almost certainly be rejected when the optimality hypothesis is true. This finding is not as problematic when the optimality restrictions are false, as we have already shown in Table 5 of section 3.2 that the unified approach has great power to reject the optimality hypothesis in that case.

[22]Refer to the appendix for a more detailed discussion of the two step procedure as well as a comparison of the estimation results with the unified approach based on GMM. The appendix also reports the results of an exercise that checks the robustness of our GMM algorithm to alternative parameter starting values.

6. Conclusions

The purpose of this paper is to demonstrate a computationally efficient method for estimating the structural parameters of various New Keynesian-style models subject to an auxiliary condition that requires the policy-rule equation minimize expected loss. Imposing an optimal-policy restriction enables joint estimation of the model parameters and the policy weights that identify central bank preferences. The empirical strategy advanced here combines the least-squares normal equations implied by the model's reduced form with the first-order necessary conditions consistent with the policymaker's control problem. The outcome is a compact set of orthogonality conditions that form the basis for estimation using GMM. In contrast to its predecessors which rely on maximum likelihood methods (e.g., Salemi (2006) and Dennis (2004)), the GMM algorithm eliminates the need to perform an optimal control exercise for each set of parameters considered during the course of estimation.

To assess the performance of our GMM approach, we conduct Monte Carlo experiments on three different New Keynesian models that differ in complexity of the structural equations and in the role of forward-looking behavior. For each model, we consider two opposing parameterizations of the policy equation. In one, the policy-rule coefficients are optimal for a given loss function, and in the other, the coefficients are not optimal for any loss function within the family that we consider. Provided the hypothesis of policy optimality is true, the Monte Carlo evidence suggests that GMM returns unbiased estimates of all structural parameters including the relative weights appearing in the central bank's objective function. Overall, the benefits from imposing optimal-policy restrictions when they are true emerge in the form of reduced uncertainty surrounding many of the key structural parameters. One shortcoming, however, is that for over-identified models, application of the standard chi-squared test rejects the optimality restrictions too often, particularly in small samples.

Perhaps our most interesting finding concerns the consequences of assuming that policy is optimal when in reality it is not. Surprisingly, Monte Carlo statistics reveal that GMM consistently delivers unbiased and precise estimates of the policy-rule coefficients regardless of whether or not the optimality hypothesis is true. In contrast, imposing false optimality restrictions tends to produce bias in some of the key structural parameters for all three models. In the course of these trials, however, application of the standard chi-squared test rejects the false optimality restrictions with very high frequency even in small samples.

We also demonstrate our GMM algorithm with an application to U.S. data. Employing data on inflation, the output gap, and the nominal interest rate, we estimate the parameters of all three New Keynesian models considered in the paper. Overall, the results suggest that the forward-looking and representative agent models fit the data well when estimation is based on least-squares normal equations alone. Augmenting the orthogonality conditions with first-order conditions from the policymaker's control problem diminishes the empirical fit of each model (especially the representative agent model) and alters estimates of some important parameters.

Acknowledgements

The authors have benefited greatly from helpful discussions with Patrick Conway, Neville Francis, Richard Froyen, Oksana Leukhina, James Stock, George Tauchen, Volker Wieland, two anonymous referees, and participants of the UNC Money-Macro seminar, the 2005 Triangle Econometrics Conference, and the 2006 Southern Economic Association annual meeting.

References

Amato, J.D., Laubach, T., 2004. Implications of habit formation for optimal monetary policy. Journal of Monetary Economics 51, 305-325.

Ball, L., 1999. Efficient rules for monetary policy. International Finance 2, 63-83.

Blancharch, O.J., Kahn, C.M., 1980. The solution of linear difference equations under rational expectations. Econometrica 48, 1305-1312.

Calvo, G.A., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12, 383-398.

Clarida, R.G., Galí, J., Gertler, M., 1999. The science of monetary policy: a new Keynesian perspective. Journal of Economic Literature 37, 1661-1707.

Clarida, R.G., Galí, J., Gertler, M., 2000. Monetary policy rules and macroeconomic stability: evidence and some theory. The Quarterly Journal of Economics 115, 147-180.

Dennis, R., 2004. Inferring policy objectives from economic outcomes. Oxford Bulletin of Economics and Statistics 66, 735-764.

Dennis, R., 2005. The policy preferences of the US Federal Reserve. Journal of Applied Econometrics 21, 55-77.

Ehrmann, M., Smets, F., 2003. Uncertain potential output: implications for monetary policy. Journal of Economic Dynamics and Control 27, 1611-1638.

Estrella, A., Fuhrer, J.C., 2002. Dynamic inconsistencies: counterfactual implications of a class of rational-expectations models. American Economic Review 92, 1013-1028.

Favero, C.A., Rovelli, R., 2003. Macroeconomic stability and the preferences of the Fed: a formal analysis, 1961-96. Journal of Money, Credit, and Banking 35, 545-556.

Fuhrer, J.C., 2000. Habit formation in consumption and its implications for monetary-policy models. American Economic Review 90, 367-390.

Fuhrer, J.C., Moore, G.R., 1995. Inflation persistence. The Quarterly Journal of Economics 110, 127-159.

Fuhrer, J.C., Rudebusch, G.D., 2004. Estimating the Euler equation for output. Journal of Monetary Economics 51, 1133-1153.

Galí, J., Gertler, M., 1999. Inflation dynamics: a structural econometric analysis. Journal of Monetary Economics 44, 195-222.

Goodfriend, M., King, R.G., 1997. The new neoclassical synthesis and the role of monetary policy, in: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual 1997. MIT Press, Cambridge, MA, pp. 231-283.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton, NJ.

Hansen, L.P., 1982. Large sample properties of generalized method of moments estimators. Econometrica 50, 1029-1054.

Heaton, J., 1995. An empirical investigation of asset pricing with temporally dependent preference specifications. Econometrica 63, 681-717.

Ilbas, P., 2007. Estimation of monetary policy preferences in a forward-looking model: a Bayesian approach. National Bank of Belgium Working Paper.

Jensen, H., 2002. Targeting nominal income growth or inflation? American Economic Review 92, 928-956.

King, R.G., Wolman, A.L., 1999. What should the monetary authority do when prices are sticky?, in: Taylor, J.B. (Ed.), Monetary Policy Rules. University of Chicago Press, Chicago, IL, pp. 349-398.

Levin, A.T., Weiland, V., Williams, J.C., 1999. Robustness of simple monetary policy rules under model uncertainty, in: Taylor, J.B. (Ed.), Monetary Policy Rules. University of Chicago Press, Chicago, IL, pp. 263-299.

Lippi, F., Neri, S., 2007. Information variables for monetary policy in an estimated structural model of the Euro area. Journal of Monetary Economics 54, 1256-1270.

McCallum, B.T., 1999. Issues in the design of monetary policy rules, in: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics. Elsevier, Amsterdam, pp. 1483-1530.

McCallum, B.T., Nelson, E., 1999a. An optimizing IS-LM specification for monetary policy and business cycle analysis. Journal of Money, Credit, and Banking 31, 296-316.

McCallum, B.T., Nelson, E., 1999b. Performance of operational policy rules in an estimated semiclassical structural model, in: Taylor, J.B. (Ed.), Monetary Policy Rules. University of Chicago Press, Chicago, IL, pp. 15-45.

Ozlale, U., 2003. Price stability vs. output stability: tales of Federal Reserve administrations. Journal of Economic Dynamics and Control 27, 1595-1610.

Rotemberg, J.J., Woodford, M., 1997. An optimization-based econometric framework for the evaluation of monetary policy, in: Bernanke, B.S., Rotemberg, J.J. (Eds.), NBER Macroeconomics Annual 1997. MIT Press, Cambridge, MA, pp. 297-346.

Rudebusch, G.D., 2001. Is the Fed too timid? Monetary policy in an uncertain world. Review of Economics and Statistics 83, 203-217.

Rudebusch, G.D., Svensson, L.E.O., 1999. Policy rules for inflation targeting, in: Taylor, J.B. (Ed.), Monetary Policy Rules. University of Chicago Press, Chicago, IL, pp. 203-246.

Salemi, M.K., 2006. Econometric policy evaluation and inverse control. Journal of Money, Credit, and Banking 38, 1737-1764.

Schmitt-Grohé, S., Uribe, M., 2007. Optimal simple and implementable monetary and fiscal policy rules. Journal of Monetary Economics 54, 1702-1725.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium model of the Euro area. Journal of the European Economic Association 1, 1123-1175.

Smets, F., Wouters, R., 2005. Comparing shocks and frictions in U.S. and Euro area business cycles: a Bayesian DSGE approach. Journal of Applied Econometrics 20, 161-183.

Söderlind, P., 1999. Solution and estimation of RE macromodels with optimal policy. European Economic Review 43, 813-823.

Söderlind, P., Söderström, U., Vredin, A., 2005. New-Keynesian models and monetary policy: a reexamination of the stylized facts. Scandinavian Journal of Economics 107, 521-546.

Svensson, L.E.O., 1997. Inflation forecast targeting: implementing and monitoring inflation targets. European Economic Review 41, 1111-1146.

Svensson, L.E.O., 1999. Inflation targeting as a monetary policy rule. Journal of Monetary Economics 43, 607-654.

Svensson, L.E.O., Woodford, M., 2005. Implementing optimal policy through inflation-forecast targeting, in: Bernanke, B.S., Woodford, M. (Eds.), The Inflation-Targeting Debate. University of Chicago Press, Chicago, IL, pp. 19-83.

Taylor, J.B., 1980. Aggregate dynamics and staggered contracts. Journal of Political Economy 88, 1-23.

Taylor, J.B., 1993. Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy 39, 195-214.

Vestin, D., 2006. Price-level versus inflation targeting. Journal of Monetary Economics 53, 1361-1376.

Walsh, C.E., 2003. Speed limit policies: the output gap and optimal monetary policy. American Economic Review 93, 265-278.

Woodford, M., 1999. Optimal monetary policy inertia. NBER Working Paper No. 7261.

Woodford, M., 2003. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton, NJ.

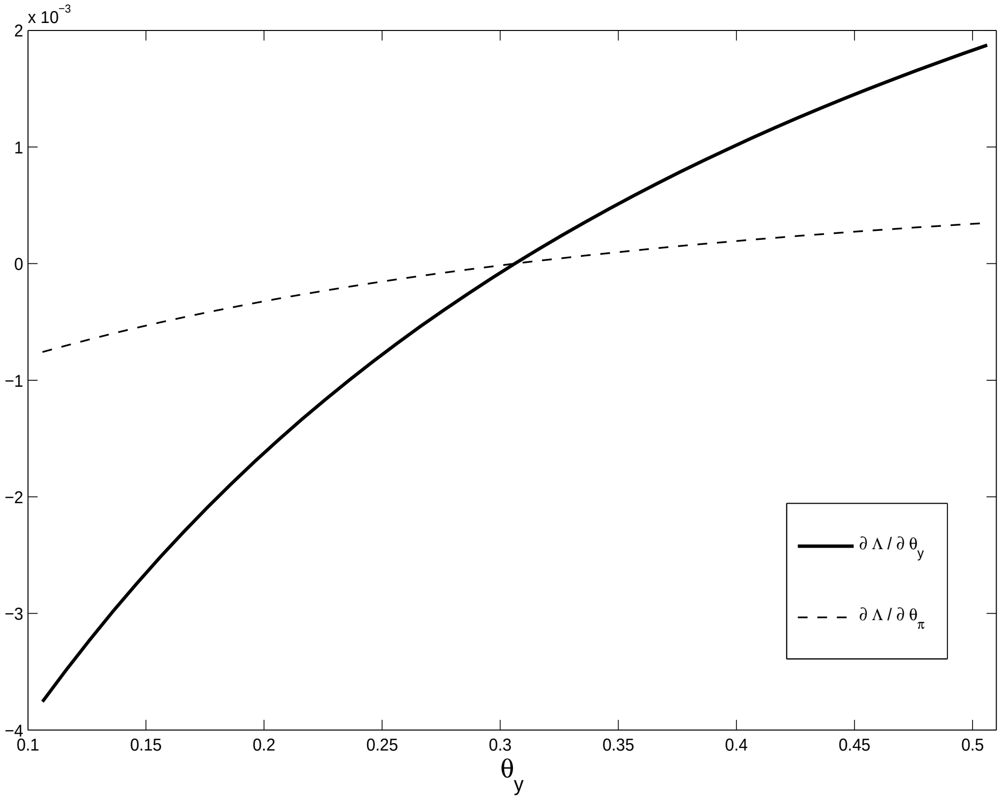

Fig. 1. The graph plots the partial derivative of the central bank loss function with respect to the parameters of the policy rule (θ_y, θ_π) for the backward-looking model. As we vary θ_y along the interval $[.1, .5]$, we hold θ_π fixed at its optimal value.

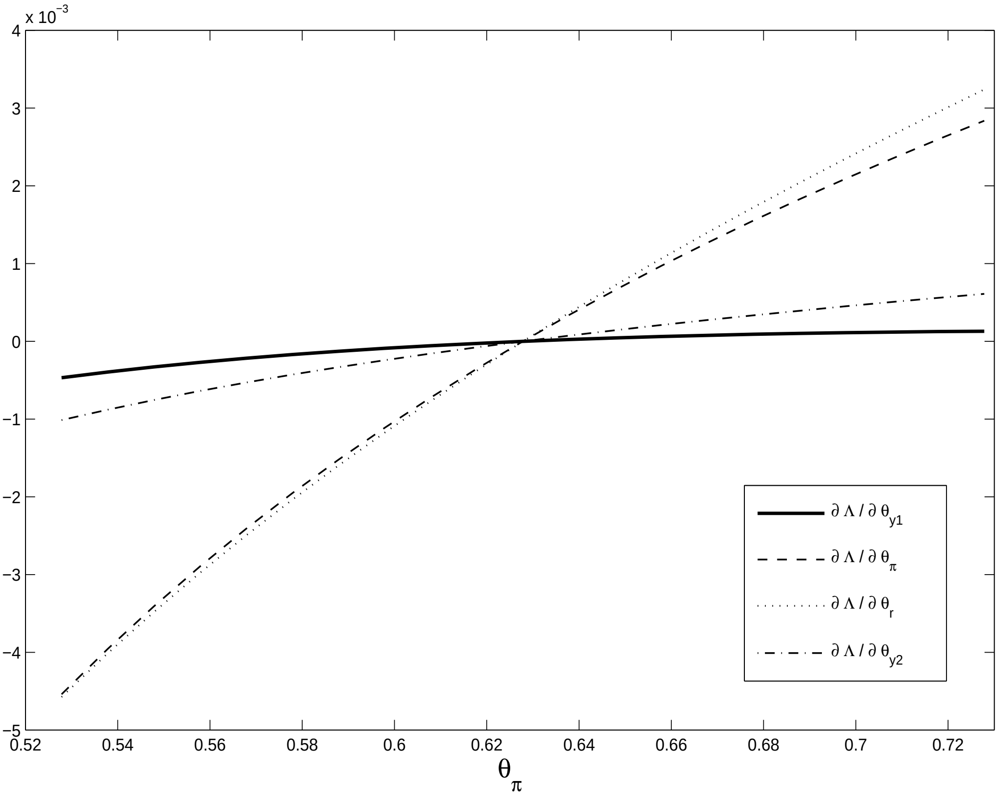

Fig. 2. The graph plots the partial derivative of the central bank loss function with respect to the parameters of the policy rule $(\theta_{y1}, \theta_{\pi}, \theta_{r}, \theta_{y2})$ for the forward-looking model. As we vary θ_{π} along the interval $[.53, .73]$, we hold θ_{y1}, θ_{r}, and θ_{y2} fixed at their optimal values.

Table 1. Backward-looking model

(ρ, θ, W)	True Value	Optimality Restriction (True): **Not Imposed** *sample size* **100**	**250**	**500**	**5000**	Optimality Restriction (True): **Imposed** *sample size* **100**	**250**	**500**	**5000**
a	0.90	0.916	0.890	0.898	0.900	0.918	0.892	0.899	0.900
		(.23)	(.04)	(.03)	(.01)	(.19)	(.04)	(.03)	(.01)
b	0.15	0.349	0.163	0.169	0.150	0.269	0.171	0.179	0.150
		(.77)	(.13)	(.09)	(.03)	(.53)	(.12)	(.09)	(.03)
α	0.50	0.495	0.493	0.494	0.500	0.497	0.493	0.496	0.500
		(.09)	(.06)	(.04)	(.01)	(.09)	(.06)	(.04)	(.01)
β	0.10	0.101	0.105	0.106	0.099	0.087	0.107	0.106	0.099
		(.08)	(.05)	(.04)	(.01)	(.07)	(.06)	(.04)	(.01)
W_y	0.10	–	–	–	–	0.066	0.131	0.129	0.111
						(.24)	(.23)	(.14)	(.05)
W_r	0.30	–	–	–	–	0.181	0.452	0.444	0.320
						(.33)	(.48)	(.41)	(.12)
θ_y	0.306	0.291	0.306	0.306	0.308	0.296	0.304	0.305	0.308
		(.09)	(.05)	(.03)	(.01)	(.08)	(.05)	(.03)	(.01)
θ_π	0.102	0.116	0.097	0.107	0.101	0.121	0.110	0.115	0.101
		(.11)	(.07)	(.04)	(.01)	(.10)	(.06)	(.04)	(.01)
Q		.24e-2	.87e-4	.62e-5	.13e-17	.64e-2	.22e-2	.69e-3	.17e-8
Fraction		1.00	1.00	1.00	1.00	0.83	0.85	0.95	1.00

1. For the case in which the hypothesis of policy optimality is **true**, the table reports estimates of the following model: $y_t = ay_{t-1} - b(r_t - \pi_t) + u_t$, $\pi_t = \alpha\pi_{t-1} + \beta y_t + v_t$, $r_t = \theta_y y_{t-1} + \theta_\pi \pi_{t-1} + w_t$. The variables are defined as: y - output, π - inflation, r - interest rate. W_y and W_r are the loss function weights for y and r normalized by the unit weight attached to π.

2. Q is the GMM estimation criterion and Fraction reports the fraction of trials that result in no outliers.

3. The parenthesis contain standard errors computed across Fraction $\times$ 100 trials for each sample size.

Table 2. Backward-looking model

(ρ, θ, W)	True Value	Optimality Restriction (False): **Imposed** *sample size* **100**	**250**	**500**	**5000**
a	0.90	0.875	0.886	0.891	0.898
		(.06)	(.04)	(.02)	(.01)
b	0.15	0.138	0.144	0.141	0.146
		(.06)	(.03)	(.02)	(.01)
α	0.50	0.485	0.483	0.485	0.487
		(.09)	(.06)	(.04)	(.01)
β	0.10	0.047	0.035	0.030	0.026
		(.05)	(.02)	(.01)	(.004)
W_y	0.10	0.4e-6	0.67e-6	0.17e-17	0.9e-18
		(.3e-5)	(.6e-5)	(.2e-17)	(.11e-17)
W_r	0.30	0.003	0.25e-2	0.22e-2	0.2e-2
		(.002)	(.07)	(.001)	(.01)
θ_y	0.20	0.177	0.182	0.184	0.186
		(.07)	(.05)	(.03)	(.01)
θ_π	2.00	2.02	2.00	2.01	2.01
		(.12)	(.07)	(.04)	(.01)
Q		0.020	0.017	0.016	0.014
Fraction		0.83	0.85	0.95	1.00

1. For the case in which the hypothesis of policy optimality is **false** and **imposed**, the table reports estimates of the following model: $y_t = ay_{t-1} - b(r_t - \pi_t) + u_t$, $\pi_t = \alpha\pi_{t-1} + \beta y_t + v_t$, $r_t = \theta_y y_{t-1} + \theta_\pi \pi_{t-1} + w_t$. The variables are defined as: y - output, π - inflation, r - interest rate. W_y and W_r are the loss function weights for y and r normalized by the unit weight attached to π.

2. Q is the GMM estimation criterion and Fraction reports the fraction of trials that result in no outliers.

3. The parenthesis contain standard errors computed across Fraction $\times$ 100 trials for each sample size.

Table 3. Forward-looking model

(ρ, θ, W)	True Value	Optimality Restriction (True): **Not Imposed** *sample size* **100**	**250**	**500**	**5000**	Optimality Restriction (True): **Imposed** *sample size* **100**	**250**	**500**	**5000**
λ	0.15	0.207	0.218	0.181	0.134	0.186	0.162	0.139	0.110
		(.25)	(.22)	(.19)	(.11)	(.23)	(.18)	(.15)	(.07)
a_1	1.10	1.04	1.04	1.07	1.12	1.02	1.05	1.08	1.14
		(.26)	(.22)	(.19)	(.11)	(.23)	(.17)	(.14)	(.07)
a_2	-0.30	-0.279	-0.271	-0.290	-0.303	-0.301	-0.300	-0.316	-0.307
		(.12)	(.08)	(.07)	(.04)	(.12)	(.08)	(.07)	(.02)
b	0.20	0.184	0.184	0.185	0.209	0.147	0.164	0.168	0.219
		(.16)	(.12)	(.09)	(.05)	(.14)	(.10)	(.08)	(.04)
α_1	0.50	0.433	0.407	0.430	0.507	0.372	0.349	0.378	0.504
		(.32)	(.24)	(.21)	(.06)	(.32)	(.23)	(.23)	(.06)
α_2	0.45	1.67	0.481	0.469	0.449	1.04	0.472	0.468	0.448
		(12.0)	(.07)	(.06)	(.02)	(4.9)	(.07)	(.06)	(.01)
β	0.15	0.196	0.187	0.180	0.150	0.185	0.191	0.184	0.150
		(.11)	(.08)	(.06)	(.02)	(.12)	(.09)	(.07)	(.01)
W_y	0.10	–	–	–	–	1.37	0.106	0.587	0.076
						(8.1)	(.22)	(3.6)	(.07)
W_r	0.30	–	–	–	–	0.749	0.209	0.228	0.314
						(6.3)	(.21)	(.19)	(.06)
θ_{y1}	1.10	1.09	1.09	1.09	1.10	1.08	1.09	1.09	1.10
		(.13)	(.09)	(.06)	(.02)	(.15)	(.11)	(.07)	(.02)
θ_π	0.63	0.628	0.610	0.625	0.627	0.646	0.635	0.642	0.627
		(.10)	(.07)	(.04)	(.02)	(.11)	(.07)	(.04)	(.01)
θ_r	0.23	0.238	0.237	0.236	0.228	0.246	0.240	0.238	0.227
		(.08)	(.05)	(.04)	(.01)	(.08)	(.04)	(.04)	(.01)
θ_{y2}	-0.20	-0.193	-0.189	-0.196	-0.197	-0.209	-0.197	-0.205	-0.197
		(.19)	(.11)	(.07)	(.02)	(.18)	(.11)	(.07)	(.02)
Q		0.032	.75e-2	.41e-2	.22e-3	0.110	0.036	0.029	.18e-2
		(.04)	(.76e-2)	(.45e-2)	(.27e-3)	(.17)	(.04)	(.05)	(.60e-2)
Fraction		1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

1. For the case in which the hypothesis of policy optimality is **true**, the table reports estimates of the following model: $y_t = \lambda E_t y_{t+1} + a_1 y_{t-1} + a_2 y_{t-2} - b(r_t - E_t \pi_{t+1}) + u_t$, $\pi_t = \beta y_t + \alpha_1 E_t \pi_{t+1} + \alpha_2 \pi_{t-1} + v_t$, $r_t = \theta_{y1} y_{t-1} + \theta_\pi \pi_{t-1} + \theta_r r_{t-1} + \theta_{y2} y_{t-2} + w_t$. The variables are defined as: y - output, π - inflation, r - interest rate. W_y and W_r are the loss function weights for y and r normalized by the unit weight attached to π.

2. Q is the GMM estimation criterion and Fraction reports the fraction of trials that result in no outliers.

3. The parenthesis contain standard errors computed across Fraction $\times$ 100 trials for each sample size.

Table 4. Forward-looking model

(ρ, θ, W)	True Value	Optimality Restriction (False): **Imposed** *sample size* **100**	**250**	**500**	**5000**
λ	0.15	0.107	0.092	0.075	0.027
		(.18)	(.13)	(.11)	(.03)
a_1	1.10	1.09	1.15	1.15	1.20
		(.19)	(.12)	(.11)	(.03)
a_2	-0.30	-0.354	-0.363	-0.361	-0.367
		(.13)	(.10)	(.06)	(.02)
b	0.20	0.181	0.202	0.206	0.239
		(.11)	(.08)	(.06)	(.02)
α_1	0.50	0.568	0.555	0.559	0.582
		(.27)	(.13)	(.13)	(.04)
α_2	0.45	0.372	0.356	0.351	0.352
		(.10)	(.08)	(.05)	(.01)
β	0.15	0.119	0.108	0.111	0.103
		(.06)	(.04)	(.03)	(.01)
W_y	0.10	0.003	0.74e-4	0.78e-4	0.3e-17
		(.02)	(.74e-3)	(.78e-3)	(.4e-17)
W_r	0.30	0.001	0.14e-3	0.46e-4	0.3e-17
		(.004)	(.07e-3)	(.46e-2)	(.5e-17)
θ_{y1}	0.50	0.489	0.479	0.468	0.471
		(.11)	(.07)	(.06)	(.01)
θ_π	1.50	1.47	1.52	1.53	1.54
		(.20)	(.09)	(.10)	(.02)
θ_r	0.50	0.512	0.509	0.506	0.506
		(.06)	(.03)	(.04)	(.01)
θ_{y2}	0.00	0.009	-0.82e-3	0.003	-0.006
		(.18)	(.09)	(.10)	(.02)
Q		0.224	0.145	0.141	0.113
		(.23)	(.09)	(.11)	(.005)
Fraction		1.00	1.00	1.00	1.00

1. For the case in which the hypothesis of policy optimality is **false** and **imposed**, the table reports estimates of the following model: $y_t = \lambda E_t y_{t+1} + a_1 y_{t-1} + a_2 y_{t-2} - b(r_t - E_t \pi_{t+1}) + u_t$, $\pi_t = \beta y_t + \alpha_1 E_t \pi_{t+1} + \alpha_2 \pi_{t-1} + v_t$, $r_t = \theta_{y1} y_{t-1} + \theta_\pi \pi_{t-1} + \theta_r r_{t-1} + \theta_{y2} y_{t-2} + w_t$. The variables are defined as: y - output, π - inflation, r - interest rate. W_y and W_r are the loss function weights for y and r normalized by the unit weight attached to π.

2. Q is the GMM estimation criterion and Fraction reports the fraction of trials that result in no outliers.

3. The parenthesis contain standard errors computed across Fraction $\times$ 100 trials for each sample size.

Table 5. Rejection frequency of over-identifying restrictions

Optimal Policy Restriction		Degrees of Freedom	Sample Size	Test Size **.25**	**.10**	**.05**	**.025**	**.01**
True	Not Imposed	1	**100**	60	40	29	17	11
			250	51	27	12	6	2
			500	52	28	17	10	5
			5000	32	13	5	2	1
True	Imposed	3	**100**	64	48	40	34	24
			250	63	43	39	32	28
			500	60	48	42	38	29
			5000	41	28	20	18	14
False	Imposed	3	**100**	98	96	96	91	81
			250	100	100	100	100	100
			500	100	100	100	100	100
			5000	100	100	100	100	100

Note: For the forward-looking model, the table reports the frequency of rejection of the over-identifying moment restrictions as a function of test size, sample size, whether or not the optimality restriction is true, and whether or not the optimality restriction is imposed during estimation. The values recorded are given in percentages and are computed across 100 trials for each sample size.

Table 6. Parameters for the representative agent model

Parameter	Description	Value
b	degree of habit formation	0.65
σ	inverse of the intertemporal elasticity of substitution	2.00
γ	degree of partial price indexation	0.75
β	household subjective discount factor	0.99*
ε	fraction of firms unable to reset prices	0.50
χ	inverse of the wage elasticity of labor supply	2.00
α	capital elasticity of output	0.33*
η	elasticity of demand for intermediate goods	11.0*
θ_π	optimal policy rule coefficient on inflation	9.28
θ_y	optimal policy rule coefficient on output	0.28
θ_r	optimal policy rule coefficient on the interest rate	1.63
W_π	implied preference weight on inflation objective	10.9**
W_y	implied preference weight on output gap objective	10.6**
δ_y	implied strength of the lag in output gap objective	0.49**

Note: * - indicates that the parameter is fixed at the given value during estimation; ** - indicates a parameter value that is implied by the values of the other parameters.

Table 7. Representative agent model

(ρ, θ, W)	True Value	Optimality Restriction (True): **Not Imposed** *sample size* **100**	**250**	**500**	**5000**	Optimality Restriction (True): **Imposed** *sample size* **100**	**250**	**500**	**5000**
b	0.65	0.702	0.670	0.669	0.650	0.739	0.698	0.688	0.650
		(.21)	(.17)	(.11)	(.03)	(.23)	(.18)	(.12)	(.03)
σ	2.00	3.34	2.75	2.19	2.04	3.32	2.28	1.98	2.02
		(4.7)	(2.4)	(1.4)	(.48)	(5.9)	(1.9)	(1.3)	(.45)
γ	0.75	0.812	1.08	0.831	0.761	0.740	0.747	0.741	0.763
		(.79)	(1.3)	(.41)	(.13)	(.41)	(.23)	(.14)	(.01)
ε	0.50	0.517	0.483	0.492	0.500	0.502	0.500	0.500	0.498
		(.10)	(.09)	(.06)	(.02)	(.04)	(.02)	(.01)	(.002)
χ	2.00	2.29	1.87	1.90	2.01	2.13	1.99	2.04	1.98
		(2.0)	(1.2)	(.80)	(.27)	(1.1)	(.58)	(.36)	(.12)
W_y	10.6	–	–	–	–	12.4	10.7	10.7	10.5
						(7.8)	(2.3)	(1.6)	(.48)
W_π	10.9	–	–	–	–	11.5	11.1	10.9	10.8
						(3.1)	(1.9)	(.95)	(.12)
θ_y	0.28	0.281	0.278	0.277	0.277	0.276	0.276	0.277	0.277
		(.06)	(.04)	(.03)	(.01)	(.06)	(.04)	(.03)	(.01)
θ_π	9.28	9.28	9.29	9.28	9.28	9.31	9.29	9.28	9.28
		(.11)	(.07)	(.04)	(.02)	(.15)	(.08)	(.05)	(.02)
θ_r	1.63	1.63	1.63	1.63	1.63	1.64	1.64	1.63	1.63
		(.02)	(.02)	(.01)	(.003)	(.03)	(.02)	(.01)	(.003)
Q		0.013	0.004	0.002	.14e-3	0.083	0.036	0.015	.78e-3
		(.01)	(.006)	(.003)	(.17e-3)	(.11)	(.06)	(.04)	(.59e-3)
Fraction		0.88	0.99	1.00	1.00	0.88	0.99	1.00	1.00

1. For the case in which the hypothesis of policy optimality is **true**, the table reports estimates of the representative agent model described in section 3.3. The parameters have the following interpretation: b - habit persistence, σ - inverse of the intertemporal elasticity of substitution, γ - partial indexation, ε - fraction of firms unable to adjust prices, χ - inverse of the wage elasticity of labor supply. $\{\theta_y, \theta_\pi, \theta_r\}$ are the coefficients of the policy rule and $\{W_y, W_\pi\}$ are the loss function weights.

2. Q is the GMM estimation criterion and Fraction reports the fraction of trials that result in no outliers.

3. The parenthesis contain standard errors computed across Fraction $\times$ 100 trials for each sample size.

Table 8. Representative agent model

(ρ, θ, W)	True Value	Optimality Restriction (False): **Imposed** *sample size* **100**	**250**	**500**	**5000**
b	0.65	0.749	0.739	0.786	0.768
		(.23)	(.21)	(.17)	(.11)
σ	2.00	2.77	3.14	1.34	0.717
		(4.4)	(5.6)	(2.2)	(.56)
γ	0.75	0.937	0.987	0.966	0.995
		(.51)	(.41)	(.32)	(.05)
ε	0.50	0.396	0.368	0.357	0.342
		(.12)	(.08)	(.07)	(.01)
χ	2.00	1.12	1.08	1.20	1.28
		(.66)	(.48)	(.39)	(.17)
W_y	10.6	10.2	10.6	7.83	6.82
		(7.2)	(10.0)	(3.7)	(.61)
W_π	10.9	8.49	5.98	5.31	4.31
		(9.4)	(4.8)	(3.7)	(.18)
θ_y	0.50	0.484	0.494	0.495	0.499
		(.08)	(.05)	(.04)	(.01)
θ_π	1.50	1.53	1.52	1.52	1.52
		(.05)	(.03)	(.02)	(.01)
θ_r	0.50	0.499	0.502	0.499	0.496
		(.03)	(.02)	(.01)	(.01)
Q		0.279	0.260	0.259	0.244
		(.14)	(.09)	(.07)	(.03)
Fraction		0.91	0.99	1.00	1.00

1. For the case in which the hypothesis of policy optimality is **false** and **imposed**, the table reports estimates of the representative agent model described in section 3.3. The parameters have the following interpretation: b - habit persistence, σ - inverse of the intertemporal elasticity of substitution, γ - partial indexation, ε - fraction of firms unable to adjust prices, χ - inverse of the wage elasticity of labor supply. $\{\theta_y, \theta_\pi, \theta_r\}$ are the coefficients of the policy rule and $\{W_y, W_\pi\}$ are the loss function weights.

2. Q is the GMM estimation criterion and Fraction reports the fraction of trials that result in no outliers.

3. The parenthesis contain standard errors computed across Fraction $\times$ 100 trials for each sample size.

Table 9. Rejection frequency of over-identifying restrictions

Optimal Policy Restriction		Degrees of Freedom	Sample Size	Test Size .25	.10	.05	.025	.01
True	Not Imposed	1	**100**	60	33	26	18	15
			250	23	10	6	5	2
			500	27	12	6	4	2
			5000	16	4	1	0	0
True	Imposed	4	**100**	36	25	19	16	16
			250	25	14	12	11	9
			500	28	9	6	4	4
			5000	19	11	6	3	2
False	Imposed	4	**100**	100	99	98	92	88
			250	100	100	100	100	100
			500	100	100	100	100	100
			5000	100	100	100	100	100

Note: For the representative agent model, the table reports the frequency of rejection of the over-identifying moment restrictions as a function of test size, sample size, whether or not the optimality restriction is true, and whether or not the optimality restriction is imposed during estimation. The values recorded are given in percentages and are computed across Fraction $\times$ 100 trials for each sample size.

Table 10. Assessment of model fit (1979:III to 2001:IV)

	Backward Looking Policy Restrictions: *Not Imposed*	*Imposed*	**Forward Looking** Policy Restrictions: *Not Imposed*	*Imposed*	**Representative Agent** Policy Restrictions: *Not Imposed*	*Imposed*
Q	4.3e-19	0.016	0.018	0.094	0.033	1.108
$Q \times T$	–	–	1.55	8.27	2.94	97.5
p-value	–	–	0.15	0.02	0.09	0.00
$\mathcal{L}$	1219.1	1215.4	1274.7	1254.7	1266.4	1083.9

Note: Q is the minimized GMM estimation criterion. $Q \times T$ is the Hansen (1982) chi-squared test statistic for the model's over-identifying restrictions. $\mathcal{L} = -\frac{T}{2}\ln(|\Phi|)$ corresponds to pseudo log likelihood and is obtained from the residual-error covariance matrix Φ.

Table 11. Structural parameter estimates (1979:III to 2001:IV)

A. Forward-Looking Model

Parameter	Optimal Policy Restrictions: *Not Imposed*	*Imposed*
λ	0.42	0.08
	(12.4)	(5.4)
a_1	0.77	1.18
	(8.9)	(.42)
a_2	-0.20	-0.30
	(3.4)	(.15)
b	0.01	0.01
	(19.7)	(3.4)
α_1	0.00	0.02
	(2.5e4)	(21.2)
α_2	0.67	0.71
	(.16)	(.36)
β	0.04	0.02
	(.74)	(1.2)

B. Representative Agent Model

Parameter	Optimal Policy Restrictions: *Not Imposed*	*Imposed*
b	0.87	0.999
	(.80)	(4.9e3)
σ	4.43	0.47e-4
	(1.6)	(5.0e3)
γ	0.76	1.00
	(.65)	(1.1e4)
ε	0.98	0.82
	(4.2)	(.32)
χ	0.02	0.9e-6
	(1.4e3)	(4.9e3)

Note: The table reports structural parameter estimates of the forward-looking and representative agent models described in sections 3.2 and 3.3. The numbers in parenthesis are estimated standard errors.

Table 12. Policy coefficient estimates (1979:III to 2001:IV)

A. Forward-Looking Model

	Optimal Policy Restrictions:			
	Not Imposed		*Imposed*	
Coefficient	Estimate	Derivative	Estimate	Derivative
θ_{y1}	0.21	-34.4	0.59	-0.9e-3
	(.18)		(.18)	
θ_{π}	0.28	-9.8	0.29	-0.2e-3
	(.11)		(.09)	
θ_{r}	0.78	6.3	0.87	-0.5e-4
	(.09)		(.08)	
θ_{y2}	-0.26	-32.7	-0.13	-0.8e-3
	(.17)		(.17)	
W_y	–	–	0.00	–
			(2.7e5)	
W_r	–	–	0.00	–
			(2.5e5)	

B. Representative Agent Model

	Optimal Policy Restrictions:			
	Not Imposed		*Imposed*	
Coefficient	Estimate	Derivative	Estimate	Derivative
θ_{y}	-0.05	-4.5e3	-1.48	0.05
	(.03)		(.96)	
θ_{π}	0.31	-7.9e2	6.69	-0.08
	(.11)		(4.5)	
θ_{r}	0.80	-17.3	-0.32e-3	-0.28
	(.08)		(.56)	
W_y	123.1	–	20.6	–
W_{π}	1.3e4	–	139.0	–

Note: The table reports policy-rule coefficient estimates and loss function weights for the forward-looking and representative agent models described in sections 3.2 and 3.3. The numbers in parenthesis are estimated standard errors.

Table 13. Comparison of GMM and FIML

(ρ, θ, W)	True Value	Optimality Restriction (True): **FIML** *sample size* **100**	**1000**	**GMM** *sample size* **100**	**1000**
λ	0.15	0.155	0.174	0.181	0.114
		(.11)	(.08)	(.23)	(.11)
a_1	1.10	0.957	0.977	1.04	1.12
		(.16)	(.11)	(.24)	(.11)
a_2	-0.30	-0.242	-0.273	-0.307	-0.315
		(.09)	(.05)	(.12)	(.04)
b	0.20	0.125	0.116	0.152	0.197
		(.08)	(.07)	(.15)	(.07)
α_1	0.50	0.453	0.461	0.372	0.448
		(.12)	(.06)	(.32)	(.17)
α_2	0.45	0.412	0.409	0.749	0.454
		(.07)	(.05)	(2.1)	(.03)
β	0.15	0.160	0.167	0.187	0.164
		(.08)	(.04)	(.13)	(.04)
W_y	0.10	0.129	0.110	12.4	0.088
		(.18)	(.07)	(101)	(.11)
W_r	0.30	0.127	0.135	0.609	0.283
		(.17)	(.12)	(4.9)	(.15)
θ_{y1}	1.10	1.06	1.07	1.08	1.10
		(.15)	(.07)	(.14)	(.04)
θ_π	0.63	0.682	0.676	0.649	0.633
		(.11)	(.06)	(.11)	(.03)
θ_r	0.23	0.239	0.242	0.244	0.233
		(.06)	(.03)	(.08)	(.02)
θ_{y2}	-0.20	-0.189	-0.195	-0.209	-0.206
		(.17)	(.08)	(.18)	(.05)

Note: For the case in which the hypothesis of policy optimality is **true** and **imposed**, the table reports estimates of the structural parameters, loss function weights, and policy-rule coefficients for the forward-looking model described in section 3.2. The estimates in the left panel are obtained from a nested approach that uses full information maximum likelihood (FIML). The estimates in the right panel are obtained using GMM. The numbers in parenthesis are estimated standard errors.

Table 14. Comparison of GMM and FIML

(ρ, θ, W)	True Value	Optimality Restriction (False): **FIML**		**GMM**	
		sample size		*sample size*	
		100	**1000**	**100**	**1000**
λ	0.15	0.151	0.144	0.107	0.045
		(.08)	(.08)	(.18)	(.05)
a_1	1.10	0.943	0.929	1.09	1.18
		(.10)	(.10)	(.19)	(.06)
a_2	-0.30	-0.196	-0.187	-0.354	-0.372
		(.08)	(.10)	(.13)	(.03)
b	0.20	0.177	0.183	0.181	0.222
		(.05)	(.05)	(.11)	(.03)
α_1	0.50	0.614	0.629	0.568	0.562
		(.07)	(.04)	(.27)	(.08)
α_2	0.45	0.461	0.445	0.372	0.347
		(.05)	(.05)	(.10)	(.04)
β	0.15	0.057	0.070	0.119	0.108
		(.02)	(.03)	(.06)	(.02)
W_y	0.10	0.008	0.007	0.003	0.7e-15
		(.3e-2)	(.3e-2)	(.02)	(.5e-14)
W_r	0.30	0.026	0.028	0.001	0.55e-4
		(.9e-2)	(.03)	(.4e-2)	(.4e-3)
θ_{y1}	0.50	0.262	0.318	0.489	0.479
		(.70)	(.52)	(.11)	(.03)
θ_π	1.50	-0.144	-0.299	1.47	1.53
		(.92)	(.70)	(.20)	(.08)
θ_r	0.50	0.639	0.649	0.512	0.510
		(.30)	(.25)	(.06)	(.02)
θ_{y2}	0.00	-0.259	-0.312	0.009	-0.010
		(.64)	(.53)	(.18)	(.06)

Note: For the case in which the hypothesis of policy optimality is **false** and **imposed**, the table reports estimates of the structural parameters, loss function weights, and policy-rule coefficients for the forward-looking model described in section 3.2. The estimates in the left panel are obtained from a nested approach that uses full information maximum likelihood (FIML). The estimates in the right panel are obtained using GMM. The numbers in parenthesis are estimated standard errors.

Table 15. Two step estimation of the forward-looking model

(ρ, θ, W)	True Value	Two Step Procedure *sample size* **100**	**1000**	**10000**	Unified Estimation *sample size* **100**	**1000**	**10000**
λ	0.15	0.207	0.157	0.131	0.181	0.114	0.112
		(.25)	(.17)	(.09)	(.23)	(.11)	(.05)
a_1	1.10	1.04	1.09	1.12	1.04	1.12	1.13
		(.26)	(.17)	(.09)	(.24)	(.11)	(.06)
a_2	-0.30	-0.279	-0.299	-0.304	-0.307	-0.315	-0.304
		(.12)	(.06)	(.03)	(.12)	(.04)	(.01)
b	0.20	0.184	0.198	0.209	0.152	0.197	0.217
		(.16)	(.08)	(.04)	(.15)	(.07)	(.03)
α_1	0.50	0.433	0.472	0.506	0.372	0.448	0.505
		(.32)	(.16)	(.04)	(.32)	(.17)	(.01)
α_2	0.45	1.67	0.459	0.449	0.749	0.454	0.450
		(12.0)	(.04)	(.01)	(2.1)	(.03)	(.01)
β	0.15	0.196	0.164	0.149	0.187	0.164	0.149
		(.11)	(.05)	(.01)	(.13)	(.04)	(.01)
W_y	0.10	0.679	0.120	0.081	12.4	0.088	0.073
		(2.7)	(.25)	(.07)	(101)	(.11)	(.06)
W_r	0.30	5.2e20	0.222	0.276	0.609	0.283	0.321
		(5.2e21)	(.14)	(.07)	(4.9)	(.15)	(.04)
θ_{y1}	1.10	1.09	1.10	1.10	1.08	1.10	1.10
		(.13)	(.04)	(.01)	(.14)	(.04)	(.01)
θ_π	0.63	0.628	0.621	0.628	0.649	0.633	0.628
		(.10)	(.03)	(.01)	(.11)	(.03)	(.01)
θ_r	0.23	0.238	0.232	0.228	0.244	0.233	0.226
		(.08)	(.03)	(.01)	(.08)	(.02)	(.01)
θ_{y2}	-0.20	-0.193	-0.199	-0.198	-0.209	-0.206	-0.196
		(.19)	(.05)	(.01)	(.18)	(.05)	(.02)

Note: The table reports estimates of the structural parameters, policy-rule coefficients, and loss function weights for the forward-looking model described in section 3.2. The figures in the left panel are obtained from a two step procedure whereby the normal equations alone are used to estimate structural parameters and policy-rule coefficients. The policymaker's first-order conditions are then used in a second step to estimate the loss function weights holding fixed the first step estimates. The figures in the right panel are obtained from the unified approach described in section 2. The numbers in parenthesis are standard errors.

Table 16. Two step estimation of the forward-looking model

Partial Derivative	Two Step Procedure *sample size*			Unified Estimation *sample size*		
	100	**1000**	**10000**	**100**	**1000**	**10000**
θ_{y1}	1.3e24	33.7	0.13e-4	-0.46e-3	-0.37e-5	0.21e-5
	(1.3e25)	(216)	(.6e-3)	(.3e-2)	(.2e-3)	(.6e-5)
θ_{π}	1.9e24	-63.6	-0.8e-3	0.21e-3	-0.28e-4	0.33e-6
	(1.9e25)	(408)	(.8e-3)	(.6e-2)	(.1e-3)	(.3e-5)
θ_{r}	3.3e25	-176	-0.13e-2	-0.53e-3	-0.58e-4	-0.35e-5
	(3.3e26)	(1.1e4)	(.2e-2)	(.2e-2)	(.1e-3)	(.9e-5)
θ_{y2}	6.1e24	238	-0.17e-3	-0.35e-3	0.68e-5	0.16e-5
	(6.1e25)	(1.6e4)	(.8e-3)	(.3e-2)	(.2e-3)	(.6e-5)

Note: The table reports estimates of the partial derivatives of loss with respect to the policy-rule coefficients of the forward-looking model described in section 3.2. The derivatives in the left panel are computed on the basis of a two step procedure. The derivatives in the right panel are computed on the basis of the unified approach to estimation. The numbers in parenthesis are standard errors.