Generalized Method of Moments and Inverse Control

Gregory Givens
Department of Economics
Jones College of Business
Middle Tennessee State University
Murfreesboro, Tennessee 37132

Michael K. Salemi
Department of Economics
CB# 3305, Gardner Hall
University of North Carolina
Chapel Hill, NC 27599-3305
Michael_Salemi@unc.edu

First Draft: May, 2005
This Draft: December, 2005

1. Introduction

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 show how to adapt GMM so that a certain type of optimality restriction is imposed in the course of estimating the parameters of a structural model.

The optimality restriction that we have in mind is that the parameters of one equation of the model are the coefficients of an optimal policy rule. For the kinds of models we consider, analytic expressions for the optimal settings of these parameters do not exist and the researcher is forced to adopt numerical strategies to impose the optimality restriction. One straightforward strategy is "brute force." With a brute force approach, the parameters of the model are estimated by quasi maximum likelihood and optimal policy-rule coefficients are found each time likelihood is computed. The brute force approach is computation intensive and the great majority of computation involves calculating optimal policies associated with model parameters very different from those that fit the data well.

In this paper, we show how to write a necessary condition for a policy rule coefficient to be optimal as a moment restriction. We then provide several examples where the structure is a macroeconomic model that determines equilibrium values of the output gap and inflation rate and optimal policy means the policy rule that minimizes a loss function. For each example, we conduct a battery on Monte Carlo experiments that assess the performance of GMM with and without the imposition of the "optimality condition" moments. Our findings, so far preliminary, indicate that our approach provides unbiased estimates that converge to the true values of the structural parameters when they should. We also find that the standard χ^2 test rejects the model too frequently when the auxiliary restrictions are true and imposed. On the other hand, we find that the standard test has great power to reject the auxiliary restrictions when they are false.

2. The Econometric Problem

Let the equations of a macroeconomic model be written as $S(X_t, \varepsilon_t, \rho, \theta) = 0$ where X_t and ε_t are (nx1) vectors of endogenous variables and serially uncorrelated exogenous shocks, and where ρ and θ are vectors of parameters. The elements of ρ are the structural parameters of the model, determined by the preferences and technology of private agents. The elements of θ are parameters of the central bank's monetary policy rule. Given ε_1 through ε_T, S determines X_1 through X_T.

One equation in S explains the behavior of the central bank. The bank sets the interest rate, r, according to a rule:

$$r_t = P(X_{t-1}, \theta) \tag{1}$$

The elements of θ determine how the bank alters the interest rate in response to economic events as represented by X_{t-1}. We assume that the bank commits to a fixed-parameter rule and chooses θ to achieve an objective that we describe later.

Given data for X, the econometrician wishes to estimate ρ and θ subject to S and the added restriction that θ satisfies the first-order conditions of the bank's optimization problem. The equations in S are structural in the sense that they are invariant across choices of θ.

For the models we consider, the reduced form implied by S is a first-order autoregression

$$X_t = G X_{t-1} + \varphi_t \tag{2}$$

where G is a (nxn) matrix of reduced-form parameters and elements of G are non-linear functions of ρ and θ and where φ_t is the (nx1) vector of reduced-form errors. Let $\Omega = [\omega_{i,j}]$ be the covariance matrix of φ_t. Since ε_t is serially uncorrelated, so is φ_t.

The objective of the central bank is to stabilize the economy in the sense of minimizing a loss function given by

$$\Lambda = E_0 \sum_{j=1}^{\infty} \delta^j (X_j - X^*_j)' W (X_j - X^*_j) \tag{3}$$

where X^*_t is the (nx1) vector of target values for X_t, W is a (nxn) matrix of weights that describe the relative importance to the central bank of stabilization objectives, and δ is the central bank's time rate of discount. The auxiliary restriction to be imposed in the course of estimation is that θ minimizes Λ.

Write Λ as $\text{trace}\ [W \cdot M]$ where $M = (1-\delta)^{-1} [\Omega + \delta G \Omega G' + \delta^2 G^2 \Omega (G^2)' + \ldots]$ and let Vector [ ] be the operator that converts an (n x n) matrix into an (n^2 x 1) vector by stacking its columns. Rewrite M as:

$$\text{Vector}[M] = (1 - \delta)^{-1} [I - \delta G \otimes G]^{-1} \text{Vector}[\Omega] \tag{4}$$

The first order conditions for an optimal policy are that the partial derivatives of Λ with respect to the elements of θ equal zero. For the k^{th} element of θ,

$$\begin{aligned} \frac{\partial \text{Vector}[M]}{\partial \theta_k} &= \frac{\delta}{1\ \delta} [I - \delta G \otimes G]^{-1} [\frac{\partial G \otimes G}{\partial \theta_k}] [I - \delta G \otimes G] \text{Vector}[\Omega] \\ &= D_k \text{Vector}[\Omega] \end{aligned} \tag{5}$$

where the terms in (5) involving the Kroneker product of G are combined in the $(n^2 \ x \ n^2)$ matrix D_k for notational convenience. The necessary conditions for an optimal policy rule are

$$(6) \qquad \frac{\partial \Lambda}{\partial \theta_k} = \text{Vector}[W]' D_k \text{Vector}[\Omega] = 0$$

Imposing Policy Optimality in the Course of Estimation

The restriction that θ minimizes Λ can be imposed by “brute force” as in Salemi (2005). The Salemi algorithm begins with initial values of ρ, computes the loss-minimizing θ for those values, solves the model for its reduced form and computes log likelihood. It searches for values of ρ that improve log likelihood and stops when no higher value can be found. The brute-force algorithm is computation intensive because it compute the loss-minimizing θ for each calculation of log likelihood.

An alternative approach is generalized method of moments (GMM). To estimate ρ and θ by GMM one combines the least squares normal equations with the moment-restriction counterpart to (6). Let $\hat{\varphi}_t$ be the (nx1) sample estimate of φ_{ϵ}. Let $\hat{\Phi}_t = [\hat{\varphi}_t \hat{\varphi}_t']$ be the (nxn) matrix of time t residual cross products. Using the sample-residual covariance matrix as an estimate of Ω, one may re-write (6) as

$$(7) \qquad \frac{\partial \Lambda}{\theta_j} = \frac{1}{T} \sum_1^T \text{Vector}[W]' D_j \text{Vector}[\hat{\Phi}_t] = 0$$

Equation (7) says that if the first order conditions hold, then a certain linear combination of the elements of the residual covariance matrix must vanish and demonstrates that the first-order equations can be written as moment restrictions.

We combine the parameters of the model into a vector (μ) and define g(μ) to be the (mx1) vector the typical element of which is the sample counterpart of a moment that should be zero if the model is true. The GMM estimation criterion is $Q = g(\mu)' S^{-1} g(\mu)$ where S^{-1} is the weighting matrix described in Hamilton (1994, p. 412-13). The following section provides several examples that use GMM to estimate ρ and θ.

3. Examples of GMM Estimation

In this section, we report what happens when we use the GMM approach described in Section 2 to estimate the parameters of three different models. The three models differ in the complexity of the relationship between structural and reduced form parameters, in the number of over-identifying restrictions they place on the reduced form, and in how the partial derivatives of the first order conditions are computed. All three models are New Keynesian in spirit. They determine the equilibrium relationship among the output gap, the inflation rate, and a short-maturity interest rate controlled by the central bank. Let y be the output gap and let p and r be the differences of inflation and the interest rate from target values. Given these transformations, it is reasonable to assume that the central bank target value for all three variables is zero.

3.1 A Backward-looking Macroeconomics Model

We begin with a backward-looking model for several reasons. First, an analytic expression for $\frac{\partial G}{\partial \theta_j}$ is available for the backward looking model and exploiting it is likely to make GMM estimation more accurate. Second, for a backward looking model, optimal values for θ may be verified independently with the matrix Ricatti equations. Third, the backward looking model we use implies exact identification of the model parameters and it turns out to be interesting to compare the results for exactly identified and over-identified models.

Our backward-looking model has three equations and is similar to that of Dennis (2002).

$$y_t = a\, y_{t-1} - b\,(r_t - p_t) + u_t \quad (8)$$
$$p_t = \alpha\, p_{t-1} + \beta\, y_t + v_t \quad (9)$$
$$r_t = \theta_y\, y_{t-1} + \theta_p\, p_{t-1} + w_t \quad (10)$$

Equation (8) is an IS schedule where output depends on its own lag and on a backward-looking measure of the real rate of interest. Equation (9) is a Phillips curve where inflation depends on its own lag and the output gap. If $\beta > 0$, then a larger output gap tends to raise inflation presumably because it implies higher marginal costs of production. Equation (10) is the policy rule that explains how the central bank changes the interest rate in response to departures of the output gap and inflation from their target values. We think of time as measured in quarters and believe it reasonable to identify the effects of policy by assuming that the central bank reacts to the state of the economy one quarter earlier. Hansen and Sargent (1980) explain why it is reasonable to include an error term in a policy rule.

We use two parameterizations of (8) - (10). The structural parameters are the same in each and imply that the solution of the model is stable for any reasonable set of policy rule coefficients. The first parameterization, reported in Table 1, includes values for the policy rule coefficients that are optimal. The second, reported in Table 2, includes policy rule coefficients that, given the structural parameters, are not optimal.

The values chosen for a and α imply persistent responses of output and inflation to shocks. The value chosen for b implies that a one percentage point increase in the interest rate lowers output by 0.15 percent, other variables unchanged. The value chosen for β implies that a one percent increase in the output gap raises the inflation rate by 0.10 percent, other variables unchanged. The values chosen for Ω are similar to those reported for quarterly U. S. data by Salemi (2005) and imply that innovations to the interest rate are positively correlated with innovations to output and inflation.

The central bank is assumed to put more weight on stabilizing inflation ($W_p = 1.0$)[1] than on stabilizing output ($W_y = 0.10$) and to place intermediate weight on stabilizing the interest rate ($W_r = 0.30$). For those weights, a policy rule with θ_y equal to 0.306 and θ_p equal to 0.102 minimizes expected policy loss.[2] Iterating the Ricatti equations to convergence and using a quasi-Newton method with a finite-difference gradient (DUMINF of the IMSL Library) to minimize expected loss produced identical values for θ_y and θ_p.

Figure 1 shows how the partial derivatives of Λ with respect to θ_y and θ_p vary with θ_y and θ_p. At the loss minimizing values, the partial derivative functions return numbers on the order of 10^{-17}. Away from the optimal values the partial derivative functions return values that range between 10^{-2} and 10^{-4}. Figure 1 suggests that a GMM estimation criterion that includes (7) can discriminate between optimal and sub-optimal values of θ_y and θ_p.

Table 2 is based on a sub-optimal policy rule where θ_y equals 0.20 and θ_p equals 2.00. While this policy rule satisfies the Taylor principle, it is not the optimal rule for any values of

[1]Since it is well known that the elements of W are identified only up to a scalar transform, we imposed $W_p = 1.0$ as a normalization.

[2]Provided that $W_r > 0$, the loss-minimizing values for both θ_y and θ_p vary with W_y and W_r. If $W_r = 0$, the loss-minimizing θ_p varies with the weights but the loss minimizing θ_y remains the same.

W_p, W_y, and W_r. Thus, the results in Table 2 show what happens to parameter estimates when a false optimality restriction is imposed on the model.

Before proceeding further, we take up an interesting issue about what moments to employ in GMM estimation. GMM is frequently based on the least squares normal equations that require that the covariance between the residuals of an equation and its regressors equal zero. For the backward-looking model there are 6 such restrictions: $E[\varphi_{jt} y_{t-1}] = E[\varphi_{jt} p_{t-1}] = 0$ for j = 1, 2 and 3. In test estimations where covariance restrictions were used without restriction (7), GMM produced unbiased parameter estimates that converged to true values as sample size increased. When (7) was added, GMM estimates frequently converged to a particular set of wrong values: b = 0.0 and a = 1.0.

To see what is going on, it is instructive to compute population values of the moments employed in GMM. Recall (2). Let G be the true reduced form coefficient matrix and H be an estimate of G. The least squares normal equations evaluated at H are:

(11) $$E[(X_t - HX_{t-1})X'_{t-1}] = (G - H)\Omega_X = 0$$

where Ω_X is the population covariance matrix of X. If (11) are the only restrictions that estimator H must satisfy, then H will tend to equal G.

For the backward model, (7) can be satisfied by two different parameterizations: the true one where b>0 and θ_y and θ_p are equal to their loss-minimizing values and a false one where b=0 and θ_y and θ_p are equal to any values. Why does the false model satisfy the first order conditions for an optimal policy? If interest rates do not affect aggregate demand (b = 0), then changes in θ_y and θ_p have no effect on expected loss. Put another way, if b = 0 any policy is as good as any other.

GMM faces a tradeoff between satisfying (7) and (11). Setting b = 0, satisfies (7) exactly but implies that (11) will not be satisfied. Setting b (and the other structural parameters) to their true values satisfies (11) but may not satisfy (7) exactly.

There is a remedy. Notice that (11) implies that the covariance between the residuals and the regressors depends not only on G -H but also on Ω_X. The same value of G - H implies a smaller value of Q, the GMM estimation criterion, if the diagonal elements of Ω_X are small than if they are large. As a remedy we remove the dependence of the least square normal equations on the scale of the data by restating the moment restrictions as correlations rather than covariances. The correlation version of the least squares normal equations was used to produce the estimates reported in this paper.[3]

To assess the performance of GMM with auxiliary moment restriction (7), we conducted a battery of Monte Carlo experiments. The hill climber used in the experiments was PATERN from Version 6 of the GQOPT Library of Fortran optimization programs (Goldfeld and Quandt, 1972). PATERN is a direct search algorithm that combines exploratory searches parallel to the parameter-space axes and "pattern" searches in directions found successful in recent iterations. Because 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. During the first set, the GMM

---

[3]To check whether our remedy was likely to work, we conducted several Monte Carlo experiments using GMM with covariance restrictions on samples generated with different innovation covariance matrices. As the innovation covariance matrix approached the identify matrix, the frequency of convergence to the false parameterization went to zero.

weighting matrix was the identity matrix. After the first set, the optimal weighting matrix was estimated with the formula given by Hamilton (1994, p. 413) and a second set of calls to PATERN was undertaken.

For the backward looking model, GMM estimation entails eight moment restrictions. Six restrictions require that the residuals of each of the three equations are uncorrelated with lagged y and p. Two restrictions require that the partial of expected loss with respect to the two policy rule coefficients vanish. Table 1 reports findings when the optimal policy hypothesis is true, Table 2 when it is false. Each Table reports findings when the optimality restriction is imposed and when it is not. Since there are six parameters to estimate when optimality is not imposed and eight when optimality is imposed, the parameters are exactly identified for both cases and both Tables. The typical cell in the tables reports parameter-estimate averages and standard deviations computed across 100 samples for several sample sizes. The data generating process for each sample is (2) where G is computed for the true values of the model and where φ_t is the output of a multi-variate normal random number generator.

There are two panels in Table 1. The first reports experiments where (7) is not imposed even though the true values of the parameters satisfy (7). For these experiments, a, b, α, β, θ_y and θ_p are estimated but W_y and W_r are not. The second panel reports experiments where (7) is imposed and all 8 parameters are estimated.

Table 1 also reports the fraction of samples where parameter estimates converged to "reasonable" values where reasonable means values of W_y and W_r less than 2.0. Fraction was always 1.0 when the optimality restriction was not imposed. When the optimality restriction was imposed, Fraction ranges from 0.83 for a sample size of 100 to 1.00 for a sample size of 1000. Faced with outliers, the econometrician would repeat estimation using different starting values or different switch settings for the hill-climber. Because it is hard to automate re-estimation, we do not include outliers in the statistics reported in Table 1.

Table 1 supports several findings. First, GMM returns estimates of a, b, α, β, θ_y and θ_p that are unbiased and converge to the true values as sample size increases whether or not the optimality restriction is imposed. Second, when the optimality restriction is imposed, GMM returns estimates of W_y and W_r that are unbiased and converge to the true values although convergence is slower than for the other parameters. Third, the average value of Q, the GMM estimation criterion, ranged between 10^{-2} for a sample size of 100 to 10^{-8} for a sample size of 5000. For the larger sample sizes, smaller values of Q were obtained when the optimality restriction was not imposed than when it was. Although they are not reported in the table, the partial derivatives of loss with respect to the policy parameters averaged 10^{-6} when sample size was 250 and 10^{-10} when sample size was 5000.

Table 2 reports the results of Monte Carlo experiments where the optimality hypothesis is false, that is where the true values of θ_y and θ_p do not minimize the loss function for any values of W_y and W_r. Since the parameters are exactly identified in the estimations underlying Table 2, we do not expect the estimation criterion (Q) to be larger in Table 2 than in Table 1. Instead, we find that the imposition of a false restriction biases the estimates of some of the model parameters. As expected, the first panel of Table 2 shows that GMM estimates of the parameters are unbiased and converge to true values with sample size. The second panel shows that no bias in estimates of a, b, α, θ_y and θ_p but large and significant bias in estimates of β. It also shows that the estimates of W_y and W_r converge to zero suggesting that the best way to reconcile the sample data with optimality is to assume that the central bank cares only about stabilizing inflation.

Overall, the results in Tables 1 and 2 suggest that one can successfully estimate the parameters of the backward looking model with the GMM algorithm that combines correlation versions of the least squares normal equations with the optimality restriction based on the partial derivative of central bank loss with respect to the parameters of the central bank interest rate rule. The results also suggest that the imposition of a false optimality restriction will bias some but not all of the parameter estimates. Notably, the false optimality restriction need not bias estimates of the coefficients of the policy rule.

3.2 A Forward-looking Macroeconomics Model

The second example is based on the forward looking macroeconomics model that Salemi (2005) uses to study econometric policy evaluation. The model is composed of an IS curve, a Phillips curve, and a rule for the short-term interest rate.

$$y_t = \lambda E_t y_{t+1} + a_1 y_{t-1} + a_2 y_{t-2} - b(r_t - E_t p_{t+1}) + u_t \tag{12}$$

$$p_t = \beta y_t + \alpha_1 E_t p_{t+1} + \alpha_2 p_{t-1} + v_t \tag{13}$$

$$r_t = \theta_y y_{t-1} + \theta_p p_{t-1} + \theta_r r_{t-1} + \theta_{y-1} y_{t-2} + w_t \tag{14}$$

Equation (12) can be obtained by combining the linearized Euler equation that characterizes a representative household's optimal choice between consumption and saving and the market clearing condition for output. As explained by Clarida, Gali, and Gertler (1999), the presence of expected future output in the IS equation results from the desire of households to smooth consumption. The presence of lagged output in the IS equation can be explained by habit persistence and delays between decision making and consumption. Woodford (1996), Rotemberg and Woodford (1997) and Bernanke, Gertler, and Gilchrist (1998) provide details.

Equation (13) is a Phillips curve. If the coefficient on lagged inflation is zero, (2) is a new Phillips curve as defined by Gali and Gertler (1999), Clarida, Gali, and Gertler (2000), Svensson (2000) and many others. If the coefficient on lagged inflation is not zero, (2) is a Phillips curve modified to account for inflation inertia. Gali and Gertler (2000), Christiano, Eichenbaum and Evans (2003), Rotemberg and Woodford (1997), and Fuhrer and Moore (1995) provide different micro-founded stories that explain the presence of lagged inflation in the Phillips curve.

Equation (14) is a policy rule that explains how the central bank sets the short term interest rate and permits the central bank to react to lagged values of the economy's state variables. As in the first example, the central bank is assumed to choose coefficients of the rule to minimize expected loss.[4]

Re-defining the state vector to be $X_t = (\, y_t \;\; p_t \;\; r_t \;\; y_{t-1}\,)'$, (12) - (14) may be written in Blanchard and Kahn (1980) format as:

$$\text{(15)} \qquad \begin{bmatrix} X_t \\ E_t\, y_{t+1} \\ E_t\, p_{t+1} \end{bmatrix} = B \begin{bmatrix} X_{t-1} \\ y_t \\ p_t \end{bmatrix} + D \begin{bmatrix} u_t \\ v_t \\ w_t \end{bmatrix}$$

where B and D are (6x6) and (6x3) matrices the elements of which are non-linear functions of the parameters of equations (12) - (14).

---

[4]In contrast to the backward looking model, the optimal policy rule coefficient for the lagged interest rate is typically positive. Given that private agents are forward looking, it is optimal for the central bank to commit to persistent changes in the rate of interest.

GMM estimation of the forward looking model is more complex than GMM estimation of the backward looking model. The reduced form of the forward looking model still is a first-order autoregression as in (2), but there is now no analytic expression for G, the matrix of reduced form coefficients. We use the method of Blanchard and Kahn (1980) to compute G.

There is no analytic expression for $\frac{\partial G}{\theta_j}$ which is needed to compute the first order condition embodied in (7). We use a symmetric finite differences to approximate $\frac{\partial G}{\theta_j}$. Figure1b shows how the four partial derivatives of Λ with respect to the parameters of the policy rule vary with departures of θ_y from its optimal value. The graphs for θ_p, θ_r, and $\theta_{y\text{-}1}$ are similar and not displayed. As with the backward looking model, the graphs suggest that an estimation criterion using (7) can discriminate between optimal and sub-optimal values of policy rule parameters.

Tables 3 and 4 present the parameterizations used for the forward looking model. In Table 3, the optimality hypothesis is true and the true values of θ_y, θ_p, θ_r, and $\theta_{y\text{-}1}$ are those than minimize loss when $W_p = 1$, $W_y = 0.10$ and $W_r = 0.30$. In Table 4, the optimality hypothesis is false and the true values of θ_y, θ_p, θ_r, and $\theta_{y\text{-}1}$ minimize loss for no values of the loss function weights.

The values chosen for the structural parameters imply that a shock to the interest rate induces a "hump-shaped" response in output, a frequently reported stylized fact of U.S. economic data. They also imply that shocks to the economy induce persistent responses in the inflation rate. And they imply that there exists a unique saddle path solution to the structural equations for a wide range of policy rules including the "do nothing" rule where the central bank simply changes the interest rate to keep pace with the inflation rate. We use the same innovation covariance matrix that we used for the backward looking model.

To assess the performance of GMM, we conducted a battery of Monte Carlo experiments using the hill climber and the algorithm described earlier. When the optimality hypothesis is not imposed, the estimation criterion is based on 12 least-squares-normal-equation restrictions. When the optimality hypothesis is imposed, estimation is based on 16 restrictions, 12 normal equation restrictions plus 4 partial derivative restrictions.

Table 3 reports findings for the case when the optimality hypothesis is true. In the first panel, the optimality hypothesis is not imposed. In the second panel it is.

Table 3 supports several findings. First, GMM returns unbiased estimates of the parameters of (12)-(14) that converge to true values with sample size whether or not restriction is imposed. Second, imposing the optimality restriction leads to more accurate estimates in the sense that the sample standard deviations of estimates are smaller when the optimality restriction is imposed than when it is not. Salemi (2005) also reports that imposing the restriction that the policy rule is optimal aids estimation of the model's other structural parameters. Third, when the optimality restriction is imposed, GMM returns estimates of W_y and W_r that are unbiased and converge to the true values of those weights as sample size increases. The partial derivatives of loss with respect to the policy rule parameters had an average value on the order of 10^{-3} when sample size was 100 and 10^{-5} when sample size was 5000.

Table 4 reports results of Monte Carlo experiments where the optimality hypothesis is false, that is where the true values of the policy rule coefficients do not minimize the loss function for any values of W_y and W_r. The results in Table 4 support several findings. First, as expected, estimates of all parameters are unbiased and converge to true values when the optimality restriction is not imposed in the course of estimation. In fact, sample standard deviations are uniformly smaller in panel one of Table 4 than in panel one of Table 5. Second, imposing the false optimality hypothesis in the course of estimation does not produce biases estimates of the policy rule coefficients. The policy rule coefficients are estimated as accurately in panel two as in panel one. Third, imposing the false optimality hypothesis does produce biases estimates of some structural parameters. The estimates of λ, the coefficient on expected future output in the IS schedule, are far below the true value. Estimates of two Phillips curve parameters are likewise too small: α_2, the coefficient on lagged inflation, and β, the coefficient of the output gap. There is also some evidence of slight bias in the other structural parameters. Fourth, estimates of W_y and W_r are near zero.

GMM estimation entails over identification of the parameters of the forward looking model. When the optimality restriction is not imposed, there are 11 parameters and 12 moment restrictions (each of the three residuals must be uncorrelated with each of the four regressors). When the optimality hypothesis is imposed, there are two additional parameters (W_y and W_r) and four additional restrictions (four partial derivatives). The standard test of the over-identifying restrictions is based on the finding that T*Q is asymptotically χ^2 with degrees of freedom equal to the number of over-identifying restrictions.

Table 5 presents the frequency of rejection of the over-identifying restrictions as a function of test size, sample size, whether or not the optimality restriction is true and whether or not it is imposed. The results in Table 5 support several findings. First, when the optimality hypothesis is not imposed, the number of rejections is too high in smaller samples but converges gradually to the expected number whether or not the optimality hypothesis is true. When the optimality hypothesis is false but is imposed, the hypothesis that the restrictions are satisfied is rejected for every sample at every test size and every sample size over 100 indicating that the test has substantial power. Third, when the optimality hypothesis is true and imposed, the standard χ^2 rejects too frequently even at large sample sizes.

Overall, the results in Tables 3 -5 support the conclusion that one can successfully estimate the parameters of the forward looking model with a GMM algorithm that combines normal equations and optimality restrictions. The results again suggest that imposing the optimality restriction when it is false biases estimates of structural parameters but not estimates of the policy rule coefficients. The standard test of over-identifying restrictions will reject the optimality hypothesis even in samples as small as 100. Unfortunately, the standard test will reject the optimality hypothesis too frequently when it is true.

3.3 A Representative Agent General Equilibrium Macroeconomics Model

The third example is a dynamic general equilibrium model in which the central bank chooses coefficients of its policy rule to maximize social welfare. In the model, a representative agent maximizes expected discounted lifetime utility subject to a standard dynamic budget constraint. Firms use labor supplied by households as the only input to a production process. They are constrained by a Calvo mechanism to change prices only in randomly chosen periods. Otherwise, firms adjust their price to keep pace with inflation. The structural equations of the model are set out in the appendix.

The model determines equilibrium values for y, p, and r. We must also keep track of y^N, the departure of output from the steady state value that would occur if prices were flexible, because maximizing social welfare amounts to minimizing a quadratic form involving the difference between y and y^N. Three exogenous shocks drive the model: ε_t is a shock to household tastes, a_t is a technology shock, and w_t is a shock to the policy instrument.

The following equations describe the movements of the four variables in the vicinity of flexible price equilibrium.

(16) $$(1+b+\beta b^2)\,y_t = b y_{t-1} + (1+\beta b+\beta b^2)E_t y_{t+1} - \beta b E_t y_{t+2} - \frac{(1-b)(1-\beta b)}{\sigma}(r_t - E_t p_{t+1}) + \frac{1-b}{\sigma}\varepsilon_t$$

(17) $$p_t - \gamma p_{t-1} = \beta E_t (p_{t+1} - \gamma p_t) + \frac{(1-\varepsilon)(1-\beta\varepsilon)}{\varepsilon} mc_t$$

(18) $$mc_t = \frac{\chi+\alpha}{1-\alpha} y_t - \frac{1+\chi}{1-\alpha} a_t - \frac{1}{1-\beta b}\varepsilon_t + \frac{\sigma}{(1-b)(1-\beta b)}\left[y_t - b y_{t-1} - \beta b E_t (y_{t+1} - b y_t)\right]$$

(19) $$r_t = \theta_p p_{t-1} + \theta_y y_{t-1} + \theta_r r_{t-1} + w_t$$

(20) $$y^N_t = \frac{1+\chi}{\chi+\alpha} a_t - \frac{\sigma(1-\alpha)}{(1-b)(1-\beta b)(\chi+\alpha)}\left[y^N_t - b y^N_{t-1} - \beta b E_t(y^N_{t+1} - b y^N_t)\right] + \frac{1-\alpha}{(1-\beta b)(\chi+\alpha)}\varepsilon_t$$

Equation (16) is the IS schedule. Habit persistence implies that the IS schedule is a two-sided distributed lag of output. Equation (16) collapses to the standard forward looking IS schedule when $b = 0$. Equation (17) is the Phillips curve. The assumption that firms index to inflation when they are not permitted to re-optimize makes inflation depend on past as well as expected future inflation. Equation (17) collapses to the standard forward looking Phillips curve when $\gamma = 0$. Equation (18) defines marginal cost. If b and γ equal zero, marginal cost is a function only of y, taste shocks, and technology shocks. When these parameters are not zero, the expression for marginal cost involves both past and expected future values of output.

Equation (19) is the policy rule. It turns out that the state variables for the model are lagged values of y, p, and r. Thus, the form of (19) permits the central bank to adjust the interest rate to all relevant information available before the current shocks are realized. Equation (20) defines the "flexible price" counterpart to the output gap and explains how it varies with productivity and output shocks. Economic behavior is governed by the ten parameters described in the following table.

Parameters of the Dynamic General Equilibrium Model		
Parameter	**Description**	
b	Strength of habit formation	0.65
σ	Inverse inter-temporal elasticity of substitution	2.00
γ	Degree of partial price indexation	0.75
β	Household discount rate	0.99*
ε	Probability that a firm may not re-optimize price in a period.	0.50
χ	Inverse of wage elasticity of labor supply	2.00
α	Capital elasticity of output	0.333*
θ	Elasticity of demand for intermediate goods.	11.0*
θ_p θ_y θ_r	Central bank policy-rule coefficients	0.276, 9.28, 1.63
*Indicates that the parameter fixed at the given value during estimation.		

Following Rotemberg and Woodford (1997), we assume that the central bank chooses a policy rule that maximizes social welfare. Social welfare is defined as the expected lifetime utility of the representative household and is approximated as the negative of the loss function

$$(21) \qquad \Lambda = \frac{1-\beta b}{1-b} E_0 \sum_{t=0}^{\infty} \beta^t \left[\lambda_p (p_t - \gamma p_{t-1})^2 + \lambda_x (x_t - \delta x_{t-1})^2 \right]$$

where x_t is defined as the difference between y_t and $y^N{}_t$. The loss function weights, λ_p and λ_x depend on the structural parameters and are defined in the appendix.

The representative-agent model presents estimation challenges that are similar in some ways to the challenges provided by the previous models. To find the reduced form again requires solution of a system of expectational difference equations. Here the reduced form is a first-order vector autoregression in y, p, and r and we again use the Blanchard and Kahn algorithm to compute the coefficients of the reduced form.

There are important differences. First, the mapping from the structural parameters to the reduced form coefficients is more complicated than for the other models because the coefficients of the expectational difference equations are themselves non-linear functions of the structural parameters. Second, the assumption that the central bank maximizes social welfare implies that the weights of the loss function are known functions of the structural parameters rather than free parameters. Estimating the model under the hypothesis that the central bank maximizes social welfare adds restrictions but no new parameters to the estimation. Third, for a given structural error covariance matrix, a change in the structural parameters implies a change in the reduced form error covariance matrix because the mapping from the structural to the reduced form errors depends on the structural parameters. It follows that the formula for the partial derivatives of loss with respect to the coefficients of the policy rule must take account of the implied changes in the reduced form error covariance matrix in order to correctly compute the first order conditions associated with an optimal policy rule.

To assess the performance of GMM, we again conducted a battery of Monte Carlo experiments. When the optimality hypothesis is not imposed, the estimation criterion is based on 9 least squares normal equation restrictions. When the optimality hypothesis is imposed, estimation is based on 12 restrictions, 9 normal equation restrictions plus 3 partial derivative restrictions.

Table 6 reports findings when the optimality hypothesis is true. The first panel reports findings when the optimality hypothesis is not imposed, the second when it is. At small sample sizes, some samples returned estimates with unreasonably large estimates of γ or σ.

Encountering this problem, a researcher would repeat estimation with different starting values or algorithm settings. Because re-starting the algorithm is difficult to automate, we delete samples that produce outliers from the statistics reported in the table.

Table 6 supports several findings. First, GMM returns unbiased estimates that converge with sample size whether or not the optimality hypothesis is imposed although the sample standard deviations of σ and ε, the two utility function elasticities, are quite large for small sample sizes. Second, policy rule coefficients are accurately and precisely estimated even in small samples whether or not the optimality restriction is imposed. Third, imposing the optimality hypothesis when it is true increases the precision with which γ, σ, and ε are estimated. Fourth, as sample size increases, the fraction of samples that produce reasonable estimates goes from a low of 88 percent at sample size 100 to 100 percent at samples sizes 500 and above. Fifth, the weights in the central bank objective function are estimated without bias. The sample standard deviations of those estimates converge quickly with sample size.

Table 7 reports results when the optimality hypothesis is false and supports the following findings. When the restriction is not imposed, GMM performs somewhat worse. At sample sizes of 100 and 250, a smaller fraction of samples produce reasonable results and sample standard deviations are larger for most of the structural parameters. For larger sample sizes, parameter estimate averages and standard deviations are close to those reported in panel one of Table 6. Second, policy rule coefficient estimates are unbiased and precisely estimated whether or not the false optimality hypothesis is imposed in the course of estimation. Third, when the optimality restriction is false and imposed, structural parameter estimates are biased and estimates of the weights of the central bank objective function are biased toward zero. As sample size increases, the weight estimates converge to values below the true values

Table 8 presents the frequency of rejection of the over-identifying restrictions as a function of test size, sample size, whether or not the optimality restriction is true and whether or not it is imposed. When the optimality restriction is not imposed, we estimate 8 parameters using 9 moment restrictions implying 1 over-identifying restriction. When it is imposed, we estimate the same 8 parameters using 12 moment restrictions implying 4 over-identifying restrictions.

When the optimality hypothesis is not imposed, Table 8 shows that the number of rejections is too high in smaller samples. For sample size 500, the number of rejections approximately equals the expected number at each test level. Surprisingly, for sample size 5000, there are too few rejections, especially when the optimality restriction is true.

When the optimality restrictions are imposed and true, the number of rejections is too large in smaller samples but converges to the expected number with sample size. When the optimality hypothesis is imposed and false, the restrictions are rejected at every test size and every sample size over 100. Even at sample size 100, the restrictions are rejected in 88 percent of the samples by a one percent test and in 98 percent of samples by a five percent test.

Overall, the results in Tables 6 through 8 suggest that one can successfully estimate the parameters of a dynamic general equilibrium model using GMM. Adding the auxiliary restrictions implied by the first order equations continues to improve estimation accuracy. However, the value-added of those restrictions is smaller for DGE than for the second model because the weights of the central bank objective function are exact functions of the structural parameters and identified whether or not the optimality restrictions are imposed. The chief value of estimation with the auxiliary restrictions is in rejecting the optimality hypothesis when it is false. The results in Table 8 show that the standard test has great power in even very small samples.

4. Conclusions

In this paper, we have demonstrated how to estimate the structural parameters of a model by GMM subject to auxiliary moment restrictions implied by the hypothesis that the policy-rule equation of the model has coefficient values that minimize a quadratic loss function. In some models, imposing the auxiliary restrictions is an example of inverse control because it permits the researcher to estimate policy objectives by observing policy actions embodied in the policy rule. The Monte Carlo experiments we conducted suggest that, if the optimal-policy restriction is true, estimation of the structural parameters is more precise when the auxiliary restrictions are imposed than when they are not. They also suggest that the relative size of the loss function weights can be estimated with reasonable precision at sample sizes as small as 250. Finally, the Monte Carlo experiments indicate that the standard chi-square test rejects the optimality hypothesis when it is false with very high frequency even when sample size is only 100.

In a dynamic general equilibrium model, one might expect the value of imposing the auxiliary restrictions to be smaller because the quadratic loss weights are exact functions of the structural parameters. Nevertheless, our Monte Carlo experiments show that imposing the optimal-policy restrictions when they are true, permits more precise estimation of several of the structural parameters. They also show that the standard test of the over-identifying restrictions implied by the optimality hypothesis has great power at even small sample sizes.

Perhaps our most interesting finding concerns what happens when the optimality hypothesis is false. Our Monte Carlo experiments show that GMM returns accurate estimates of the policy rule coefficients whether or not the optimality hypothesis is true and whether or not the optimality restrictions are imposed in the course of estimation. They suggest, for all three of the models we consider, that imposing the optimality restrictions when the hypothesis is false produces bias not in the policy-rule coefficients but in the structural parameters.

In future work, we intend to take our estimation strategy to the data and to expand the class of models that we consider.

Appendix

The appendix is under construction.

Literature Cited

Atoian, Rouben V., Gregory E. Givens, and Michael K. Salemi, "Policy evaluation with a forward-looking model," in Minford, Patrick, ed., *Money Matters: Essays in Honour of Alan Walters*, 2003, Edward Elgar, Cheltenham, U.K., 294-316.

Bernanke, Ben, Mark Gertler, and Simon Gilchrist, "The financial accelerator in a quantitative business cycle framework," NBER Working Paper 6455, March, 1998.

Blanchard, Olivier J. and Charles M. Kahn "The solution of linear difference models under rational expectations," *Econometrica* 48, 1980, 1305-11.

Calvo, Guillermo A., "Staggered prices in a utility-maximizing framework," *Journal of Monetary Economics* 12, 1983, 383-98.

Christiano, Lawrence J., Martin Eichenbaum, and Charles L. Evans, "Nominal Rigidies and the Dynamic Effects of a Shock to Monetary Policy," Revision of NBER Working Paper 8403, 2003.

Clarida, Richard, Jordi Gali, and Mark Gertler, The science of monetary policy: a new Keynesian perspective," *Journal of Economic Literature* XXXVII, 1999, 1661-1707.

_______________________, "Monetary policy rules and macroeconomic stability: Evidence and some theory," *The Quarterly Journal of Economics*, 115, 2000, 147-80.

Dennis, Richard, "The policy preferences of the US Federal Reserve," Federal Reserve Bank of San Francisco mimeo, April, 2002.

Fuhrer, Jeffrey C. and George R. Moore, "Monetary policy trade-offs and the correlation between nominal interest rates and real output," *American Economic Review* 85, 1995a, 219-239.

______________, "Inflation persistence," *Quarterly Journal of Economics*, February, 1995b, 129-59.

Gali, Jordi and Mark Gertler, "Inflation dynamics: a structural econometric analysis," *Journal of Monetary Economics* 44, 1999, 195-222.

Givens, Gregory E., "Revisiting the Delegation Problem in a Sticky Price and Wage Model," University of North Carolina mimeo, 2004.

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

Hansen, Lars Peter, "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 4, 1982, 1029-54.

Hansen, Lars P. and Thomas J. Sargent, “Formulating and estimating dynamic linear rational expectations models,” *Journal of Economic Dynamics and Control* 2, 1980, 7-46.

Klein, Paul, “Using the generalized Schur form to solve a multivariate linear rational expectations model,” *Journal of Economic Dynamics and Control 24*, 2000, 1405-23.

Rotemberg, Julio J. and Michael Woodford, “An optimization-based econometric framework for the evaluation of monetary policy,” in Bernanke, Ben S. and Julio J. Rotemberg, eds., NBER Macroeconomics Annual 1997, MIT Press, Cambridge, MA, 1997, 297-346.

Sack, Brian and Volker Weiland, “Interest-rate smoothing and optimal monetary policy: A review of recent empirical evidence,” *Journal of Economics and Business*, 52, 2000, 205-28.

Salemi, Michael K., “Revealed preference of the Federal Reserve: using inverse-control theory to interpret the policy equation of a vector autoregression,” *Journal of Business and Economic Statistics* 13, 1995, 419-433.

_____________, “Econometric Policy Evaluation and Inverse Control,” forthcoming in the Journal of Money Credit and Banking, 2005.

Svensson, Lars E. O., “Open-economy inflation targeting,” *Journal of International Economics*, 50, 2000, 155-83.

Taylor, John B., “Estimation and Control of a Macroeconomic Model with Rational Expectations,” *Econometrica 47*, 1979, 1267-86.

Woodford, Michael, “Control of the public debt: A requirement for price stability?,” NBER Working Paper 5684, 1996.

_______________, “Interest and Prices.” Princeton University Working Paper, 2002

_______________, “Optimal Monetary Policy Inertia,” NBER Working Paper No 7261, 1999.

Table 1
Sample Average and Standard Deviation of GMM Parameter Estimates
Backward-Looking Model [a]
Optimality Hypothesis is True

Parms	True Values	Optimality Restriction is Not Imposed				Optimality Restriction is Imposed			
		Number of Observations in Sample				Number of Observations in Sample			
		100	250	500	5000	100	250	500	5000
a	0.90	.916 (.23)	0.890 (.04)	0.898 (.03)	0.900 (.01)	.918 (.19)	0.892 (.04)	0.899 (.03)	0.900 (.01)
b	0.15	.349 (.77)	0.163 (.13)	0.169 (.09)	0.150 (.03)	.269 (.53)	0.171 (.12)	0.179 (.09)	0.150 (.03)
α	0.50	.495 (.09)	0.493 (.06)	0.494 (.04)	0.500 (.01)	.497 (.09)	0.493 (.06)	0.496 (.04)	0.500 (.01)
β	0.10	.101 (.08)	0105 (.05)	0.106 (.04)	0.099 (.01)	.087 (.07)	0.107 (.06)	0.106 (.04)	0.099 (.01)
W_y	0.10	NA	NA	NA	NA	.066 (.24)	0.131 (.23)	0.129 (.14)	0.111 (.05)
W_r	0.30	NA	NA	NA	NA	.181 (.33)	0.452 (.48)	0.444 (.41)	0.320 (.12)
θ_y	0.306	.291 (.09)	0.306 (.05)	0.306 (.03)	0.308 (.01)	.296 (.08)	0.304 (.05)	0.305 (.03)	0.308 (.01)
θ_p	0.102	.116 (.11)	0.097 (.07)	.107 (.04)	0.101 (.01)	.121 (.10)	0.110 (.06)	0.115 (.04)	0.101 (.01)
Q [b]		.24e-02	.87e-04	.62e-05	0.13e-17	.643e-02	0.22e-02	.69e-03	.17e-08
Fraction [c]			1.00	1.00	1.00	.83	.85	0.95	1.00

a. The model is:

$$y_t = a\,y_{t-1} - b\,(r_t - p_t) + u_t$$
$$p_t = \alpha\,p_{t-1} + \beta\,y_t + v_t$$
$$r_t = \theta_y\,y_{t-1} + \theta_p\,p_{t-1} + w_t$$

where y is the output gap, p is the inflation rate, and r is the interest rate. The first equation is an IS schedule, the second a Phillips curve, and the third a monetary policy rule. W_y and W_r are the central-bank-loss-function weights for y and r. The weight for p is normalized to 1.0

b. Q is the minimized GMM estimation criterion.

c. The fraction of samples not resulting in outlying estimates of W_y or W_r where estimates greater than 2.0 are considered to be outliers.

Table 2
Sample Average and Standard Deviation of GMM Parameter Estimates
Backward-Looking Model[a]
Optimality Hypothesis is False

Parms	True Values	Optimality Restriction is Not Imposed				Optimality Restriction is Imposed			
		Number of Observations in Sample				Number of Observations in Sample			
		100	250	500	5000	100	250	500	5000
a	0.90	.885 (.06)	.892 (.04)	.896 (.03)	.901 (.01)	.875 (.06)	.886 (.04)	.891 (.02)	.898 (.01)
b	0.15	.149 (.06)	.152 (.03)	.147 (.02)	.150 (.01)	.138 (.06)	.144 (.03)	.141 (.02)	.146 (.01)
α	0.50	.498 (.09)	.497 (.06)	.497 (.04)	.499 (.01)	.485 (.09)	.483 (.06)	.485 (.04)	.487 (.01)
β	0.10	.113 (.08)	.109 (.05)	.105 (.03)	.099 (.01)	.047 (.05)	.035 (.02)	.030 (.01)	.026 (.004)
W_y	0.10	NA	NA	NA	NA	.4e-06 (.3e-05)	.67e-06 (.6e-05)	.17e-17 (.2e-17)	.90e-18 (.11e-17)
W_r	0.30	NA	NA	NA	NA	.0026 (.002)	.248e-02 (.07)	.220e-02 (.1e-02)	.20e-02 (.01)
θ_y	0.20	.198 (.08)	.200 (.05)	.200 (.03)	.202 (.01)	.177 (.07)	.182 (.05)	.184 (.03)	.186 (.01)
θ_p	2.00	2.01 (.11)	1.99 (.07)	2.01 (.04)	2.00 (.01)	2.02 (.12)	2.00 (.07)	2.01 (.04)	2.01 (.01)
Q[b]		.31e-02	.61e-18	.58e-18	.56e-18	.020	.017	.016	.014

a. The model is:

$$y_t = a\,y_{t-1} - b\,(r_t - p_t) + u_t$$
$$p_t = \alpha\,p_{t-1} + \beta\,y_t + v_t$$
$$r_t = \theta_y\,y_{t-1} + \theta_p\,p_{t-1} + w_t$$

where y is the output gap, p is the inflation rate, and r is the interest rate. The first equation is an IS schedule, the second a Phillips curve, and the third a monetary policy rule. W_y and W_r are the central-bank-loss-function weights for y and r. The weight for p is normalized to 1.0

b. Q is the minimized GMM estimation criterion.

Table 3 Sample Average and Standard Deviation of GMM Parameter Estimates Forward-Looking Model [a] Optimality Hypothesis is True									
Parms	**True Values**	**Optimality Restriction is Not Imposed**				**Optimality Restriction is Imposed**			
		Sample Size				**Sample Size**			
		100	**250**	**500**	**5000**	**100**	**250**	**500**	**5000**
λ	0.15	0.207 (.25)	0.218 (.22)	0.181 (.19)	0.134 (.11)	0.186 (.23)	0.162 (.18)	0.139 (.15)	0.110 (0.07)
a_1	1.10	1.04 (.26)	1.04 (.22)	1.07 (.19)	1.12 (.11)	1.02 (.23)	1.05 (.17)	1.08 (.14)	1.14 (.07)
a_2	-0.30	-0.279 (.12)	-0.271 (.08)	-0.290 (.07)	-0.303 (.04)	-0.301 (.12)	-0.300 (.08)	-0.316 (.07)	-0.307 (.02)
b	0.20	0.184 (.16)	0.184 (.12)	0.185 (.09)	0.209 (.05)	0.147 (.14)	0.164 (.10)	.168 (.08)	0.219 (.04)
α_1	0.50	0.433 (.32)	0.407 (.24)	0.430 (.21)	0.507 (.06)	0.372 (.32)	0.349 (.23)	0.378 (.23)	0.504 (.06)
α_2	0.45	1.67 (12.0)	0.481 (.07)	0.469 (.06)	0.449 (.02)	1.04 (4.9)	0.472 (.07)	0.468 (.06)	0.448 (.01)
β	0.150	0.196 (.11)	0.187 (.08)	0.180 (.06)	0.150 (.02)	0.185 (.12)	0.191 (.09)	0.184 (.07)	0.150 (.01)
W_y	0.10	NA	NA	NA	NA	1.37 (8.1)	0.106 (.22)	0.587 (3.6)	0.076 (.07)
W_r	0.30	NA	NA	NA	NA	0.749 (6.3)	0.209 (.21)	0.228 (.19)	0.314 (.06)
θ_y	1.10	1.09 (.13)	1.09 (.09)	1.09 (.06)	1.10 (.02)	1.08 (.15)	1.09 (.11)	1.09 (.07)	1.10 (.02)
θ_p	0.628	0.628 (.10)	0.610 (.07)	0.625 (.04)	0.627 (.02)	0.646 (.11)	0.635 (.07)	0.642 (.04)	0.627 (.01)
θ_r	0.228	0.238 (.08)	0.237 (.05)	0.236 (.04)	0.228 (.01)	0.246 (.08)	0.240 (.04)	0.238 (.04)	0.227 (.01)
θ_{y-1}	-0.20	-0.193 (.19)	-0.189 (.11)	-0.196 (.07)	-0.197 (.02)	-0.209 (.18)	-0.197 (.11)	-0.205 (.07)	-0.197 (.02)
Q		0.0318 (.04)	0.753e-02 (.76e-02)	.407e-02 (.45e-02)	0.22e-03 (0.27e-03)	0.110 (.17)	0.036 (.04)	0.0285 (.05)	0.178e-02 (.60e-02)

a) The model is:

$$y_t = \lambda E_t y_{t+1} + a_1 y_{t-1} + a_2 y_{t-2} - b(r_t - E_t p_{t+1}) + u_t$$
$$p_t = \beta y_t + \alpha_1 E_t p_{t+1} + \alpha_2 p_{t-1} + v_t$$
$$r_t = \theta_y y_{t-1} + \theta_p p_{t-1} + \theta_r r_{t-1} + \theta_{y-1} y_{t-2} + w_t$$

where y is the output gap, p is the inflation rate, and r is the interest rate. The first equation is an IS schedule, the second a Phillips curve, and the third a monetary policy rule. W_y and W_r are the central-bank-loss-function weights for y and r. The weight for p is normalized to 1.0

b) GMM is the minimized GMM estimation criterion.

Table 4
Sample Average and Standard Deviation of GMM Parameter Estimates
Forward-Looking Model[a]
Optimality Hypothesis is False

Parms	True Values	Optimality Restriction is not Imposed				Optimality Restriction is Imposed			
		Sample Size				Sample Size			
		100	250	500	5000	100	250	500	5000
λ	0.15	0.122 (.17)	0.133 (.14)	0.133 (.13)	0.142 (.05)	0.107 (.18)	0.092 (.13)	0.075 (.11)	0.027 (0.03)
a_1	1.10	1.10 (0.15)	1.11 (.11)	1.10 (.11)	1.10 (.04)	1.09 (.19)	1.15 (.12)	1.15 (.11)	1.20 (.03)
a_2	-0.30	-0.332 (.14)	-0.312 (.10)	-0.308 (.06)	-0.299 (.01)	-0.354 (.13)	-0.363 (.10)	-0.361 (.06)	-0.367 (.02)
b	0.20	0.192 (.10)	0.201 (.08)	0.200 (.07)	0.204 (.03)	0.181 (.11)	0.202 (.08)	.206 (.06)	0.239 (.02)
α_1	0.50	0.500 (.23)	0.485 (.13)	0.486 (.12)	0.503 (.03)	0.568 (.27)	0.555 (.13)	0.559 (.13)	0.582 (.04)
α_2	0.45	0.457 (.09)	0.47 (.06)	0.445 (.04)	0.451 (.01)	0.372 (.10)	0.356 (.08)	0.351 (.05)	0.352 (.01)
β	0.150	0.150 (.07)	0.154 (.05)	0.157 (.03)	0.150 (.01)	0.119 (.06)	0.108 (.04)	0.111 (.03)	0.103 (.01)
W_y	0.10	NA	NA	NA	NA	0.0031 (.02)	0.74e-04 (..74e-03)	0.78e-04 (.78e-03)	0.3e-17 (.4e-17)
W_r	0.30	NA	NA	NA	NA	0.0013 (.004)	0.14e-03 (.07e-03)	0.46e-04 (.46e-02)	0.3e-17 (.5e-17)
θ_y	0.50	0.497 (.11)	0.491 (.07)	0.49 (.06)	0.498 (.02)	0.489 (.11)	0.479 (.07)	0.468 (.06)	0.471 (.01)
θ_p	1.50	1.49 (.12)	1.49 (.08)	1.50 (.06)	1.50 (.01)	1.47 (.20)	1.52 (.09)	1.53 (.10)	1.54 (.02)
θ_r	0.50	0.505 (.06)	0.494 (.04)	0.497 (.04)	0.500 (.01)	0.512 (.06)	0.509 (.03)	0.506 (.04)	0.506 (.01)
θ_{y-1}	0.00	0.011 (.16)	0.0234 (.09)	0.0169 (.08)	0.004 (.02)	0.0088 (.18)	-.82e-03 (.09)	0.0028 (.10)	-0.0056 (.02)
Q		.0866 (.18)	0.0340 (.011)	.0227 (.10)	0.193e-03 (.28e-03)	0.224 (.23)	0.145 (.09)	0.141 (.11)	0.113 (.005)

a) The model is:

$$y_t = \lambda E_t y_{t+1} + a_1 y_{t-1} + a_2 y_{t-2} - b(r_t - E_t p_{t+1}) + u_t$$
$$p_t = \beta y_t + \alpha_1 E_t p_{t+1} + \alpha_2 p_{t-1} + v_t$$
$$r_t = \theta_y y_{t-1} + \theta_p p_{t-1} + \theta_r r_{t-1} + \theta_{y-1} y_{t-2} + w_t$$

where y is the output gap, p is the inflation rate, and r is the interest rate. The first equation is an IS schedule, the second a Phillips curve, and the third a monetary policy rule. W_y and W_r are the central-bank-loss-function weights for y and r. The weight for p is normalized to 1.0

b) GMM is the minimized GMM estimation criterion.

Table 5
Frequency of Rejections of Over-identifying Restrictions
Forward-Looking Model

Optimal Policy Restriction		Degrees of Freedom	Sample Size	Test Size				
True	Imposed			.25	.10	.05	.025	.01
Yes	No	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
No	No	1	100	56	36	22	19	17
			250	37	19	15	11	8
			500	33	17	11	5	4
			5000	24	12	4	2	1
Yes	Yes	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
No	Yes	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

Table 6
Sample Average and Standard Deviation of GMM Parameter Estimates
Dynamic General Equilibrium Model [a]
Optimality Hypothesis is True

Parms	True Values	Optimality Restriction is Not Imposed				Optimality Restriction is Imposed			
		Sample Size				Sample Size			
		100	250	500	5000	100	250	500	5000
b	0.65	0.702 (.21)	0.670 (.17)	0.669 (.11)	0.650 (.03)	0.739 (.23)	0.698 (.18)	0.688 (.12)	0.65 (.03)
σ	2.00	3.34 (4.7)	2.75 (2.4)	2.19 (1.4)	2.04 (.48)	3.32 (5.9)	2.28 (1.9)	1.98 (1.3)	2.02 (.45)
γ	0.75	0.812 (.79)	1.08 (1.3)	0.831 (.41)	0.761 (.13)	0.740 (.41)	0.747 (.23)	0.741 (.14)	0.763 (.01)
ε	0.50	0.517 (.10)	0.483 (.09)	0.492 (.06)	0.500 (.02)	0.502 (.04)	0.500 (.02)	0.500 (.01)	0.498 (.002)
χ	2.00	2.288 (1.99)	1.87 (1.2)	1.90 (.80)	2.01 (.27)	2.132 (1.1)	1.99 (.58)	2.04 (.36)	1.98 (.12)
W_y	10.6	--	--	--	--	12.4 (7.8)	10.7 (2.3)	10.7 (1.6)	10.5 (.48)
W_p	10.9	--	--	--	--	11.5 (3.1)	11.1 (1.9)	10.9 (0.95)	10.8 (.12)
θ_y	0.276	0.281 (.06)	0.278 (.04)	0.277 (.03)	0.277 (.01)	0.276 (.06)	0.276 (.04)	0.277 (.03)	0.277 (.01)
θ_p	9.28	9.28 (.11)	9.29 (.07)	9.28 (.04)	9.28 (.02)	9.31 (.15)	9.29 (.08)	9.28 (.05)	9.28 (.02)
θ_r	1.63	1.63 (.02)	1.63 (.02)	1.63 (.01)	1.63 (.003)	1.64 (.03)	1.64 (.02)	1.63 (.01)	1.63 (.003)
Q		.0126 (.01)	.0042 (.006)	.0022 (.003)	.138e-03 (.17e-03)	.083 (.11)	.036 (.06)	.015 (.04)	.775e-03 (.59e-03)
Fraction		.88	.99	1.00	1.00	.88	.99	1.00	1.00

a) The model is the dynamic general equilibrium model described in the text and appendix where b governs habit persistence, σ is the inverse inter-temporal elasticity of substitution, γ is the degree of partial price indexation, ε is the probability that a firm cannot change price in the current period, χ is the inverse elasticity of labor supply , the θ's are the coefficients of the policy rule, and the W's are the weights of the loss-function approximation to social welfare.

b) GMM is the minimized GMM estimation criterion.

c) The fraction of samples for which the algorithm converged to reasonable estimates.

Table 7
Sample Average and Standard Deviation of GMM Parameter Estimates
Dynamic General Equilibrium Model [a]
Optimality Hypothesis is False

Parms	True Values	Optimality Restriction is Not Imposed				Optimality Restriction is Imposed			
		Sample Size				Sample Size			
		100	250	500	5000	100	250	500	5000
b	0.65	0.713 (.22)	0.667 (.16)	0.667 (.09)	0.650 (.03)	0.749 (.23)	0.739 (.21)	0.786 (.17)	0.768 (.11)
σ	2.00	3.91 (7.5)	3.24 (4.2)	2.30 (1.8)	2.05 (.52)	2.77 (4.4)	3.14 (5.6)	1.34 (2.2)	0.717 (.56)
γ	0.75	0.615 (.69)	1.23 (1.7)	1.14 (1.5)	0.780 (.19)	0.937 (.51)	0.987 (.41)	0.966 (.32)	0.995 (.05)
ε	0.50	0.580 (.14)	0.493 (.15)	0.485 (.13)	0.498 (.04)	0.396 (.12)	0.368 (.08)	0.357 (.07)	0.342 (.01)
χ	2.00	2.86 (4.2)	2.05 (1.0)	1.97 (.63)	2.01 (.20)	1.12 (.66)	1.08 (.48)	1.20 (.39)	1.28 (.17)
W_y	10.6	--	--	--	--	10.2 (7.2)	10.6 (10.0)	7.83 (3.7)	6.82 (.61)
W_p	10.9	--	--	--	--	8.49 (9.4)	5.98 (4.8)	5.31 (3.7)	4.31 (.18)
θ_y	0.50	0.509 (.06)	0.503 (.04)	0.500 (.03)	0.501 (.01)	0.484 (.08)	0.494 (.05)	0.495 (.04)	0.499 (.01)
θ_p	1.50	1.50 (.05)	1.50 (.03)	1.50 (.02)	1.50 (.006)	1.53 (.05)	1.52 (.03)	1.52 (.02)	1.52 (.01)
θ_r	0.50	0.501 (.03)	0.504 (.02)	0.501 (.01)	0.500 (.004)	0.499 (.03)	.502 (.02)	0.499 (.01)	0.496 (.01)
Q		.0140 (.01)	.0043 (.006)	.236e-02 (.039e-02)	.177e-03 (.28e-03)	.279 (.14)	.260 (.09)	.259 (.07)	.244 (.03)
Fraction		.69	.90	1.00	1.00	.91	.99	1.00	

a) The model is the dynamic general equilibrium model described in the text and appendix where b governs habit persistence, σ is the inverse inter-temporal elasticity of substitution, γ is the degree of partial price indexation, ε is the probability that a firm cannot change price in the current period, χ is the inverse elasticity of labor supply , the θ's are the coefficients of the policy rule, and the W's are the weights of the loss-function approximation to social welfare.

b) GMM is the minimized GMM estimation criterion.

c) The fraction of samples for which the algorithm converged to reasonable estimates.

Table 8
Frequency of Rejections of Over-identifying Restrictions
DGE Model

Optimal Policy Restriction		Degrees of Freedom	Sample Size	Test Size				
True	**Imposed**			**.25**	**.10**	**.05**	**.025**	**.01**
Yes	**No**	**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
No	**No**	**1**	100	64	49	35	25	17
			250	26	9	6	3	1
			500	22	12	8	6	2
			5000	17	12	5	3	0
Yes	**Yes**	**3**	100	36	25	19	16	16
			250	25	14	12	11	9
			500	28	9	6	4	4
			5000	19	11	6	3	2
No	**Yes**	**3**	100	100	99	98	92	88
			250	100	100	100	100	100
			500	100	100	100	100	100
			5000	100	100	100	100	100

Figure 1
Partial Derivatives of the Loss Function with Respect to Parameters of the Policy Rule
Figure 1a–Backward Looking Model

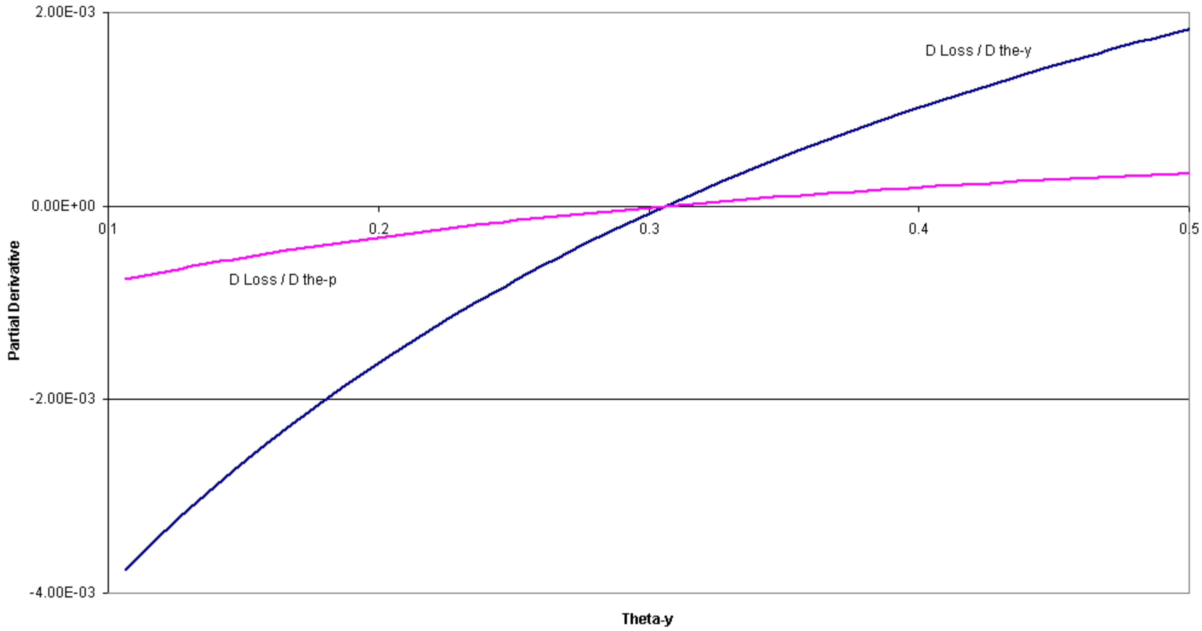

Figure 1b–Forward Looking Model