Estimating the Euler Equation for Output[#]

Jeffrey C. Fuhrer,[*,a]   Glenn D. Rudebusch[b]

[a] Federal Reserve Bank of Boston, 600 Atlantic Ave., Boston, MA 02106, USA
[b] Federal Reserve Bank of San Francisco, 101 Market Street, San Francisco, CA 94105, USA

October 2003

Abstract

New Keynesian macroeconomic models have generally emphasized that expectations of future output are a key factor in determining current output. The theoretical motivation for such forward-looking behavior relies on a straightforward generalization of the well-known Euler equation for consumption. In this paper, we use maximum likelihood and generalized method of moments (GMM) methods to explore the empirical importance of output expectations. We find little evidence that rational expectations of future output help determine current output, especially after taking into account the small-sample bias in GMM.

[*] Corresponding author: Tel. + 1-617-973-3410; fax: + 1-617-619-8410

E-mail addresses: Jeff.Fuhrer@bos.frb.org (J.C. Fuhrer); Glenn.Rudebusch@sf.frb.org (G.D. Rudebusch)

1. Introduction

A central question in macroeconomic modeling is the extent to which private agents' behavior should be modeled as "forward-looking." Standard theory implies that agents place a substantial weight on expected future income and returns in determining their consumption and investment, and on expected future prices in setting current prices. However, the weight on the future that seems consistent with both the aggregate time series data and theoretical models has been the subject of much debate. This issue is perhaps most salient in estimates of the "New Keynesian" (or "new synthesis" or "optimizing IS-LM") model, which has been a very popular framework for macroeconomic analysis recently.[1] In its simplest form, the New Keynesian model is composed of an output equation, which generalizes the consumption Euler equation to the whole economy, and an inflation equation, which resembles an expectations-augmented Phillips curve. The inflation equation has been the subject of much empirical investigation, which, again, has focused on the key issue of the relative importance of expected future inflation versus past inflation in determining current prices.[2]

In contrast, estimates of the New Keynesian output equation have been extremely rare, and there are no comprehensive empirical assessments of the relative importance of past versus future output for the determination of current output. Rotemberg and Woodford (1999) are among the first to estimate a very simple output Euler equation in which there is no weight on past inflation. However, Fuhrer (2000) and Estrella and Fuhrer (2002) show that this Euler equation, which is based on the simplest model of optimizing household behavior, provides a remarkably poor fit to the time series data on aggregate output. Fuhrer (2000) obtains much better empirical results by enhancing the model of consumer behavior with a habit formation process that adds significant inertial output dynamics. Unfortunately, as noted in Rudebusch (2002), it is not easy to infer from Fuhrer's (2000) model the exact weight on expected future output in a single equation for overall aggregate demand, and there are essentially no other available estimates with quarterly data.[3]

This gap in the literature is quite serious. New Keynesian models are often used for monetary

---

[1] For example, Svensson (1999), Clarida, Gali, and Gertler (1999), McCallum and Nelson (1999), and Rotemberg and Woodford (1999).

[2] For a survey and an analysis of these results, see Rudebusch (2002) and Rudd and Whelan (2003).

[3] Direct estimates are obtained in Smets (2000) using GMM with synthetic data for the euro-area at an annual frequency. For the U.S., also see Dennis (2003).

policy analysis to give recommendations regarding optimal monetary policy. To have confidence in these recommendations, it is important that the link between the short-term policy interest rate and aggregate demand be modeled correctly in the output equation. Accordingly, this paper provides empirical estimates from quarterly U.S. data of an Euler equation for aggregate output with a particular focus on the relative weight on future and lagged output.

The next section describes our empirical specification. It is a hybrid equation used by many authors to allow for both expectational and inertial dynamics. In order to ensure robustness, we examine several variants of the basic equation that include alternative definitions of the output gap, the real interest rate, and the information set used in forming expectations. Section 3 provides the resulting empirical estimates using maximum likelihood (ML) and generalized method of moments (GMM) procedures. Across all specifications, our ML estimates indicate a significant inertial element for output determination with the weight on future output always less than 0.5 and quite often insignificantly different from zero. The GMM estimates also indicate that output exhibits a significant inertial component across all specifications; however, they differ from the ML estimates by suggesting a significant forward-looking element as well.

The somewhat disparate nature of our results echoes earlier research on the New Keynesian inflation equation, where GMM estimates also suggested more "forward-looking" behavior than ML estimates (for example, Gali and Gertler 1999 and Fuhrer 1997). We examine this apparent dependence of the results on the estimation methodology in Section 4 with a Monte Carlo simulation experiment to determine the extent to which the ML and GMM estimators may exhibit small-sample bias. The experiment finds that the GMM estimator biases the coefficient on future output toward 0.5. That is, GMM overestimates the weight on future output when it is in fact low (below 0.5) and underestimates the weight when it is high. The ML estimator, in contrast, is unbiased. These results provide strong evidence for viewing the GMM estimates with some suspicion.

2. The Output Euler Equation

Much of the appeal of the New Keynesian model lies in its dynamic general equilibrium foundations, as derived from a neoclassical model augmented to include temporary nominal price

rigidities.[4] In the simplest case, output is determined by the intertemporal Euler equation:

$$y_t \;=\; E_t y_{t+1} \;-\; \sigma(i_t \;-\; E_t \pi_{t+1}) \;+\; \eta_t, \tag{1}$$

where y_t is the output gap (detrended output), $E_t y_{t+1}$ is the expectation formed at time t of future output at time $t + 1$, i_t is the one-period nominal interest rate, $E_t \pi_{t+1}$ is the expectation of future inflation, and η_t represents an aggregate demand shock. Of course, this equation is valid only in an economy without capital, durable goods investment, foreign trade, or a government. In this case, output equals consumption (which is the economy's aggregate resource constraint), and the path of output reflects optimal consumption choices. With the appropriate functional form for the underlying utility function, the parameter σ can then be interpreted as the intertemporal elasticity of substitution.

However, as a number of authors have demonstrated (e.g., Cogley and Nason 1995 and Estrella and Fuhrer 2002), simple descriptions of output such as (1) have a difficult time matching key dynamic features of the aggregate data. As a result, "hybrid" models that generalize (1) have been developed (e.g., Fuhrer 2000, McCallum and Nelson 1999, Rudebusch 2002, Svensson 1999, and Dennis 2003). In this spirit, our empirical investigation considers an expanded output Euler equation of the form,

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \mu E_{t-\tau} y_{t+1} - \beta E_{t-\tau}[\frac{1}{\kappa}\Sigma_{j=0}^{\kappa-1}\left(i_{t+j+m} - \pi_{t+j+m+1}\right)] + \eta_t, \tag{2}$$

which generalizes the simple Euler equation in four ways: the addition of two lags of output, the flexible timing of when expectations are formed (with the parameter τ controlling the timing), the possible influence of lagged real rates (for example, the time $t-1$ real rate will matter when $m = -1$), and the use of (potentially) longer-term interest rates with the parameter κ setting the duration of the ex ante real rate. Note that equation (2) nests the simple Euler equation (1) when $\alpha_1 = \alpha_2 = \tau = m = 0$ and $\mu = \kappa = 1$.

All of these generalizations of the basic Euler equation may be quite important. First, it is widely agreed that some adjustment process must be added to the model in order to match the inertial responses of output that are apparent in the data. As a practical matter, of course, the assumption that all output is consumed ignores sectors that are widely thought to be important sources of business cycle dynamics. However, even when applied to the time series data on

---

[4]For general derivations and discussion, see Koenig (1993), Woodford (1996), Goodfriend and King (1997), Walsh (1998), Clarida, Gali, and Gertler (1999), and Svensson (1999).

just aggregate consumption spending for nondurable goods and services (which is about 60 percent of total output), Fuhrer (2000) finds that the simple Euler equation requires significant modification to fit the data. Specifically, he augments the model of consumer behavior to include a habit formation process, in which consumers' utility depends partly on current consumption relative to past consumption. This process builds in an inertial response necessary to match the consumption data. The remainder of output other than nondurable and services consumption appears to exhibit an even greater inertial response, and various theoretical approaches have been considered in the literature. An Euler equation for investment is derived and discussed in Oliner, Rudebusch, and Sichel (1995, 1996), where it is shown that the inertial response from including adjustment costs and time-to-build lags is necessary to help match the data. Casares and McCallum (2000) come to a broadly similar conclusion. In order to capture such inertia in a parsimonious way, we add two lags of output to the canonical equation.

A second generalization in equation (2) allows for lags in the information set used to form expectations. In the standard theory, as in (1), time t expectations are used because consumption decisions at time t are determined by contemporaneous output and real interest rate expectations (the case when $\tau = 0$). However, with quarterly data, the possibility of a lag of a few months in gathering and processing information implies that expectations formed at time $t - 1$ ($\tau = 1$) may be a more appropriate specification.[5] These timing delays may affect the estimated relative weights on expected and lagged output, perhaps boosting the importance of forward-looking behavior, so we provide estimates with $\tau = 0$ and $\tau = 1$.

Finally, our hybrid output equation allows for a flexible definition of the relevant interest rate for output. The simple Euler equation (1) assumes that the 1-quarter interest rate affects the economy by inducing intertemporal substitution in consumption. Much evidence, however, suggests that purchases of durable goods, such as housing and capital, are the most interest-rate sensitive elements of aggregate demand and that long-term interest rates play a key role in this sensitivity. Varying κ in (2) controls the duration of the interest rate. In addition, varying m allows for some decision lag. Both of these factors are contained in more traditional specifications of aggregate demand such as the one in Fuhrer and Moore (1995), which is also nested in (2) when $\mu = 0$, $\kappa = 40$, $\tau = 0$, and $m = -1$.

---

[5] Such expectational timing lags have been common in New Keynesian inflation equations (see Rudebusch 2002 and Mankiw and Reis 2001). Also, in their output equation, Rotemberg and Woodford (1999) go even further and assume a two-quarter delay ($\tau = 2$), an option we do not consider.

For our investigation, the duration of the interest rate is particularly important because there can be a close relationship between the importance of forward-looking output expectations and the term of the interest rate. To see the connection, iterate equation (1) forward to obtain the "solved-forward" version of the Euler equation

$$y_t = -\sigma E_{t-\tau}[\Sigma_{j=0}^{\infty} (i_{t+j} - \pi_{t+1+j})] + \eta_t, \tag{3}$$

in which output depends only on the expected sum of future short-term real interest rates, which may approximate a long-term real rate.

3. Evidence on the Importance of Expectations

3.1. Data and Estimation Methodology

The data employed are quarterly from 1966Q1 to 2000Q4.[6] The interest rate, i_t, is the quarterly average of the overnight federal funds rate, expressed in units of percent per year. Inflation is the annualized log change in the price index. Denoting the GDP chain-weighted price index by P_t, inflation is defined as $\pi_t = 400(\ln P_t - \ln P_{t-1})$. The output gap, y_t, is 100 times the log difference between real GDP and a measure of potential output. Given the uncertainty surrounding the proper detrending of output, we use a variety of measures of potential output in order to ensure the robustness of the results.[7] The results presented below consider five different definitions of potential output: (1) the Hodrick-Prescott (HP) filter of log real GDP; (2) a one-sided band pass (BP) filter of log real GDP; (3) a segmented deterministic linear trend for log real GDP, with a breakpoint in the first quarter of 1974; (4) a deterministic quadratic trend for log real GDP; and (5) the Congressional Budget Office's (CBO) official estimate of potential output.[8]

---

[6] The beginning of the estimation period is dictated by the behavior of monetary policy. Only after 1965 did the federal funds rate, the interest rate in our study, exceed the discount rate and hence act as the primary instrument of monetary policy. Because the output Euler equation hinges on the interaction between output and interest rates, this consideration is important.

[7] Although each of these detrending methods has been used by many researchers, it is worth noting that each represents a rather imperfect attempt to overcome our inability to jointly model the determinants of long-run and business cycle behavior

[8] The HP filter is perhaps the most widely used, but it is essentially a two-sided moving average that may blur inference about the relative importance of past and future. The BP filter is the one-sided "random walk" filter described in Christiano and Fitzgerald (2003).

We estimate the output Euler equation using maximum likelihood (ML) and GMM procedures, which are described in detail in the appendix. In brief, the ML estimation is conducted by solving for the expectations of output, inflation, and the funds rate that are consistent with the rational expectations solution of the structural output Euler equation (2) and two unrestricted vector autoregressive (VAR) equations for inflation and the funds rate, which represent the dynamics of the rest of the economy in a general, agnostic fashion.[9] Then, the likelihood of the solved model can be computed for any set of parameters under the assumption that the innovations in the model are joint normally distributed with mean zero. Finally, a sequential quadratic programming algorithm is used to find the set of parameters that maximize the value of the likelihood function. Asymptotic standard errors are computed using the Berndt-Hall-Hall-Hausman (BHHH) method.

The alternative method of estimation is GMM, which essentially instruments for the expectational terms without imposing any of the model structure. The GMM estimation is straightforward because equation (2) is linear in variables and parameters. The only issue of interest is obtaining a good instrument set, and we present results for two basic instrument sets. The first comprises four lags of the funds rate, inflation, and the output gap. The second attempts to construct plausibly exogenous instruments, and includes four lags of real defense expenditures, relative oil prices, and the political party of the sitting President of the United States. We also consider alternative timing of these instrument sets as described below.

3.2. Output Euler Equation Estimation Results

Tables 1 and 2 present results for ML and GMM estimation, respectively. As indicated in the four model specification columns, both tables consider many of the number of model variations described above. These specifications differ in the output trend procedure (namely, the BP or HP filter, a segmented linear trend, labeled "Seg.", a quadratic trend, labeled "Quad.", or using the CBO's potential output), the timing of expectations (τ), and the timing (m) and duration (κ)

---

[9]The VAR equations that "close" the system contain four lags, though results were robust to variation. The VAR coefficients are estimated by OLS (separately for each definition of the output gap) and held fixed in the ML estimation. The Euler equation estimates and standard errors were essentially unchanged in a full system estimation in which the VAR parameters were jointly estimated with the Euler equation parameters. Alternatively, Dennis (2003) obtains results consistent with ours by including with the output equation other structural equations.

of the real interest rate. The remaining columns provide selected coefficient estimates, standard errors, and p-values for certain hypothesis tests.[10]

In the ML estimates, the coefficients on the lagged output terms (α_1 and α_2) are highly significant in economic and statistical terms for every specification. In economic terms, the estimates of the sum $\alpha_1 + \alpha_2$, which vary from .53 to .99 in Table 1, all indicate substantial persistence in output. In every case, the hypothesis that these coefficients are both zero can be rejected with overwhelming confidence. Indeed, because these results are so uniform across specifications (the p-values are zero to several decimal places in each case), they are not reported in the table. The clear implication is that lags of output in this specification are absolutely essential.

In contrast, the ML results on the importance of expectations are decidedly mixed. The median ML μ estimate is 0.36, but the distribution of estimates is bimodal, with one cluster of estimates at 0.0 and a second around 0.40. We examine likelihood ratio tests for the hypotheses $\mu = 1$ and $\mu = 0$. The $\mu = 1$ hypothesis can be rejected with overwhelming confidence in all cases, so we do not include the results of that test in the table. In contrast, the hypothesis that $\mu = 0$ is rejected in less than half of the cases (15 of 33 cases at the 10 percent significance level or better). Although there is not a perfect mapping, the standard errors for μ are generally consistent with the LR tests, in that high estimates of μ with small standard errors generally reject the test for $\mu = 0$. Still, a number of the standard errors imply an implausibly high degree of precision for the key parameters that is not always reflected in the p-values for the LR tests. We will explore this feature more fully in the Monte Carlo exercises below.[11]

However, it is important to note that many of the cases in which we obtain larger estimates of μ entail imprecise or economically negligible estimates of the link between output and real interest rates (the parameter β). Indeed, there is a clear negative relationship between the estimated size of the expectational parameter μ and the size of the interest rate sensitivity parameter β. For example, for the segmented trend case with $\kappa = 4$ and $\tau = m = 0$, the

---

[10] For the estimates with $\kappa = 40$, we use an approximation to the 40-quarter ex ante real interest rate (as in Fuhrer and Moore, 1995); namely, with $D = 40$: $r_t = \frac{D}{1+D} E_t r_{t+1} + \frac{1}{1+D}(i_t - E_t \pi_{t+1})$.

[11] In particular, some of the high μ estimates for the HP-filtered output gap appear many standard deviations from 0, and yet the likelihood ratio test is not able to reject $\mu = 0$. This may indicate that the standard errors for μ, which assume a quadratic approximation to the likelihood surface, are not completely reliable for this case. Alternatively, it may indicate mis-specification of the basic model.

estimate of μ is 0.43 and significantly different from zero, but the hypothesis that $\beta = 0$ cannot be rejected. Of course, when the interest rate channel is removed from this simple model, its macroeconomic content for monetary policy is eviscerated. Even in cases with a large and significant μ estimate and in which the hypothesis of $\beta = 0$ can be rejected, the size of $\hat{\beta}$ is so low as to raise questions about its economic significance. Specifically, a relatively small $\hat{\beta}$ implies that the monetary "transmission" channel is very weak and that monetary policy would have difficulty in controlling the economy.

The expectation viewpoint date assumption has no material effect on the estimates of μ or β. Table 1 provides results for $\tau = 1$ for the HP output gap definition, but omits the remaining results for compactness.

Overall, the ML results suggest that the appropriate estimate of the forward-looking component in this simple aggregate demand specification is quite small, and in many cases is not significantly different from zero. Because the linchpin of the specification for monetary policy is the economic significance of the real interest rate, the results also tend to favor a specification in which the real rate appears with a lag (whether we use a one-period or a multi-period real rate). In these cases, the dominant estimate of the forward-looking parameter is zero.

The GMM estimates in Table 2 provide an interesting contrast to the ML estimates. The model specification is identical, although rather than explicitly denoting the value of τ, it is implicit in the dating of the instrument set in column 2. For the $\tau = 0$ case, we use the exogenous instruments with lags of 0 to 4 quarters, and for the $\tau = 1$ case, we use instruments with lags of 1 to 4 quarters. We also report the p-values for the J-statistic, which, with only two exceptions, are never significant at conventional levels. Thus the instruments we have chosen, both lagged "endogenous" variables and purportedly "exogenous" variables, generally appear valid. That is, one cannot reject the orthogonality of the instruments with respect to the error term η_t.[12] The GMM results are remarkably consistent across instrument sets, real rate definitions, the timing of the real rate, and output gap proxies, so we do not present all permutations in Table 2.

As in the ML estimates, the coefficients on the lagged output terms (α_1 and α_2) are highly significant in economic and statistical terms across all specifications. Although the sum of the GMM estimates of α_1 and α_2, which is generally in the 0.45 to 0.55 range, is always lower than for the ML estimates, lagged output remains essential in the specification. Conversely,

---

[12] Therefore, the instruments satisfy one of two necessary instrument conditions: orthogonality. As for the other condition, we present results on instrument "relevance" below.

the GMM estimates of μ are uniformly larger than the ML estimates, with the median GMM estimate equal to 0.56. As expected, the less efficient GMM estimator finds uniformly larger standard errors for both μ and β than the ML estimator.

Still, just as in Table 1, the high estimates of μ are accompanied by very low estimates of β. In fact, the real interest rate coefficient estimates are not only economically minute but often of the wrong sign. All in all, the GMM estimates are not encouraging for this simple specification. While the GMM estimates center on a considerably larger forward-looking component than the ML estimates, the associated GMM real rate coefficients are not significantly different from zero. The absence of any significant real rate effect casts serious doubt on whether the GMM estimator is able to identify an "optimizing IS" or "aggregate demand" or "output Euler" equation.

In the following section, we examine the finite-sample properties of the GMM and ML estimators using a simple Monte Carlo simulation exercise in order to reconcile Tables 1 and 2.

4. Investigating ML and GMM Estimators

In our empirical results, the GMM estimates of μ lie almost uniformly above the ML estimates. There are two ways in which these estimates could be reconciled. First, one (or both) of the estimators may be asymptotically inconsistent, which might cause divergent empirical results even in small samples. Second, although consistent, one (or both) of the estimators may display small sample biases that push the GMM estimates of μ above the ML ones.

One potential concern with the consistency of the ML estimates is that they rely on an underlying assumption of the normality of the structural shocks. However, in many applications, notably in finance, the distribution of conditional errors may exhibit fat tails or other departures from normality. Thus, we considered a battery of normality tests for the maximum likelihood residuals. Table 3 shows the results for a representative specification (namely, using the CBO estimate of potential output with $\tau = m = 0$ and $\kappa = 1$). Using the Lilliefors, Cramer-von Mises, Watson, and Anderson-Darling tests, we fail to reject normality of the Euler equation and inflation (VAR) equation residuals at conventional significance levels. The reduced-form error for the funds rate (VAR) equation rejects, but sub-sample estimates reveal that this is due solely to the period 1979-1982, during which the Federal Reserve switched to a nonborrowed reserves operating procedure, allowing the federal funds rate to fluctuate dramatically. Before and after this period, we cannot reject normality. Thus we feel reasonably confident that the

normality assumption required for our ML estimator is plausible.

Of course, there may be other, more basic misspecifications in our model that could affect the ML and GMM estimates in different ways and account for the different estimates of degree of forward-lookingness. Indeed, Jondeau and Le Bihan (2003) argue that certain kinds of model misspecification such as omitted dynamics could cause the ML estimator to overstate μ and the GMM estimator to understate μ–even asymptotically (see Hall and Inoue 2003). However, the direction and size of such biases depend on the exact nature of the model misspecification hypothesized. Instead, in what follows, we focus on the small-sample behavior of the ML and GMM estimators under the assumption that the model is correctly specified. We do this in large part because the potential for small-sample bias in the GMM estimator has been documented in many contexts (see, for example, Fuhrer, Moore, and Schuh 1995, West and Wilcox 1994, and Stock, Wright, and Yogo 2002).

4.1. Baseline Monte Carlo results

In order to examine the finite-sample properties of both estimators, we design a simple Monte Carlo experiment. We assume that a simplified version of equation 2 defines y_t, given definitions for i_t and π_t:

$$y_t = (1 - \mu)y_{t-1} + \mu E_t y_{t+1} - \beta(i_t - E_t \pi_{t+1}) + \varepsilon_t. \tag{4}$$

This equation is augmented by VAR equations for the nominal interest rate and inflation, which use parameters and an error covariance matrix estimated from the data. We set β to 0.5 throughout, and examine estimates of both β and μ when μ takes the true values $[0.1, 0.5, 0.9]$.

We use this data-generating process to compute 5000 replications of simulated data for various sample sizes, with shocks drawn from a multivariate normal distribution with a covariance matrix as estimated from U.S. data, 1966-2000.[13] We then estimate the equation above using GMM, with three lags of y, i, and π as instruments. These comprise the complete set of valid instruments (given the true model structure), and given the imposed iid structure of the error terms, they should be exogenous.[14] We then estimate the same model using ML, with the same

[13]The process is essentially the reverse of the ML estimation procedure outlined above. Given a sequence of shocks and initial conditions for y, i, and π, we can generate sequences of realizations for these three variables that are consistent with the shocks, the structure of the model, and the rational expectations assumption. Details are provided in the appendix.

[14]This instrument set comprises the set of valid instruments in that, under the assumed data-generating process,

VAR equations closing the model. The estimation methods for GMM and ML are as described in the appendix.

Table 4 summarizes the results of the exercise. The top panel reports summary statistics for estimates of μ. The true value, μ^T, is shown in the third column, followed by the mean of the estimates, the median, and the cumulative density evaluated at the true value, $F(\mu^T)$. As a benchmark case, the distribution for a median unbiased estimator would (by definition) have half its mass on either side of the true estimate, or $F(\mu^T) = 0.5$.

A few key results emerge in panel A. First, the ML estimates are always centered on the true parameter value even with a sample size of only 125 observations. Second, the GMM estimates exhibit a distinct finite-sample bias. At a sample size of 125, GMM overstates the true value by about 0.2 when the true value is 0.1, and understates the true value by about 0.15 when the true value is 0.9. That is, the GMM estimator biases the estimate towards 0.5 from either side of 0.5.[15] The bias is smaller for large samples; however, the median GMM parameter estimate becomes as accurate as ML only when the sample size increases to 5000. Not surprisingly, the median standard error for the GMM estimates of μ, shown in the final column of Table 4, is two to three times larger than that for the ML estimates.

Panel B in Table 4 displays the corresponding estimates of β, the interest rate elasticity. In all cases, the true value of β is 0.5. GMM generally provides a downwardly biased estimate of β when the true value of μ is small. For example, for a sample of 125 and $\mu = 0.1$, GMM biases the estimate of β downward by about 0.2. The bias is small to nonexistent for larger values of μ, and it disappears fairly rapidly as sample size is increased. Even for the sample size of 125, the ML estimator provides accurate estimates of β regardless of the value of μ. Again, the standard errors from the GMM estimator are roughly twice as large as the standard errors from the ML estimator.

Overall, these results suggest that one should place considerably greater faith in the estimates from the ML estimator, which appears to behave quite reliably in the relevant sample size and

---

every realization of the endogenous variables may be expressed as a specific linear combination of these lags of the endogenous variables.

[15] A more general conclusion might be that GMM biases the estimate towards OLS. Because the output gap is quite autocorrelated for all measures studied here, OLS might be expected to place roughly equal weight on last quarter's and next quarter's output gap. Of course, in the case of highly correlated regressors, one might expect the standard errors to be considerably larger as well, but this implication is not evident in our exercises.

across a range of values of μ. The ML estimator exhibits no bias in the estimation of either of the two critical parameters in the specification, whereas the GMM estimator provides biased estimates of both.

4.2. Why Does GMM Fail?

It is of some interest to determine the source of the GMM estimator's poor finite-sample performance. The leading candidate for explaining the differences between GMM and ML estimators is instrument relevance or "weak" instruments (see Nelson and Startz 1990a,b; Hall, Rudebusch and Wilcox 1996; Fuhrer, Moore and Schuh 1995; Shea 1997; Staiger and Stock 1997; Stock, Wright and Yogo 2002; Mavroeidis 2003; and Stock and Yogo 2003). A sizable literature has documented the finite-sample problems that can arise if instruments do not explain sufficient variation in the instrumented variables in the first-stage regressions. Less well known are possible problems that arise in dynamic models when the rational expectations restrictions that are implied by the model are not imposed in constructing instruments. Our ML estimator imposes these restrictions, and as we will show in the next sub-section, using the optimal instruments implicitly computed by ML appears to solve the problem of weak instruments that arises with the conventional GMM estimator.

Table 5 presents instrument relevance tests for a subset of the various model specifications that we have considered. (Similar results were obtained for the other cases.) The first two columns of results present simple F-statistics from first-stage regressions for expected output and the real interest rate (denoted F-stat(y) and F-stat(ρ)), using the instrument sets indicated. This simple test provides no evidence of weak instruments in these datasets. The F-statistics are quite large–with a median value of about 36. In contrast, Staiger and Stock (1997) suggest declaring instruments to be weak if the first-stage F-statistic is less than ten. This suggests that the instruments capture a significant portion of the variation in the key endogenous variables. However, strictly speaking, Staiger and Stock's suggestion only applied to the case of a single endogenous regressor because, as noted by Shea (1997) and Stock and Yogo (2003), the first-stage F-statistic considers the explanatory power for each endogenous regressor in isolation from the others. For valid inference in our model, we require a strong set of instruments for both expected output and the real interest rate simultaneously, but the usual first-stage F-statistic does not provide information on this joint condition.

To incorporate information on this joint condition, the third and fourth columns of results in Table 5 present the R^2 statistics for regressions of the orthogonalized regressors on orthogonalized instruments, the so-called "partial R^2" recommended by Shea (1997). The $R^2(y)$ column indicates the explanatory power of the instrument set for the output gap once the instruments are orthogonalized relative to their contribution to explaining the real interest rate. Similarly, the $R^2(\rho)$ column indicates the "partial" fit for the real interest rate. These statistics suggest little or no evidence of an instrument relevance problem, as the median partial R^2 value is 0.56.

Unfortunately, the partial R^2 statistic, although intuitive, is relatively ad hoc and not well grounded for formal statistical inference. Stock and Yogo (2003) develop a more rigorous alternative. They compute critical values for a test based on Donald and Cragg's (1993) multivariate version of the F-statistic. For tractability, they consider a very conservative test, namely, whether the "poorest" or worst-behaved linear combination of the instruments provides sufficient information about the included endogenous variables in the IV regression. The test reported in the final column of Table 5 is the minimum eigenvalue of the generalized F-test, which Stock and Yogo denote G_T, where

$$G_T = \hat{\Sigma}_{VV}^{-1/2\prime} Y^{\perp\prime} P_{Z^\perp} Y^\perp \hat{\Sigma}_{VV}^{-1/2}/K_2 \tag{5}$$

and

$$\hat{\Sigma}_{VV} = Y' M_{\underline{Z}} Y/(T - K_1 - K_2) \tag{6}$$

and Y is the Txn matrix of included endogenous variables, $\underline{Z}$ is the $Tx(K_1 + K_2)$ matrix of included and excluded exogenous variables $\overset{\frown}{X \vdots Z}$, the $\perp$ denotes orthogonalization with respect to X, the TxK_1 matrix of included exogenous variables, and the P_i and M_i matrices denote projection and orthogonalization, respectively, with respect to variable i.

The Stock-Yogo test results in Table 5 provide an interesting contrast to the earlier results, although it is worth noting that the selection of critical values is not straightforward. To assess the null hypothesis that the instruments are weak, Stock and Yogo (2003) provide asymptotic critical values that depend on n and K_2 and on the maximum acceptable bias of the IV estimator relative to OLS. If one is willing to accept a bias as high as 20 percent, then the appropriate critical value is 6.2, and all but two of the cases in Table 5 would fail to reject the hypothesis of weak instruments. If a bias of 10 percent is acceptable, then the critical value is 12.8, and in only one case can we reject weak instruments.

The battery of tests that we employ provides mixed evidence on the strength of the instrument set for this problem. Still, we suspect that the univariate F-statistics are unduly optimistic, as they do not take into account the requirement that the instrument set simultaneously explain both included endogenous variables. The partial R^2 statistics should provide a better indication of the relevance of the instrument set, and seem to provide evidence that the instruments are adequate. However, no distribution theory is available for these statistics, so it is difficult to know whether a partial R^2 of 0.5 is large enough to avoid small-sample relevance problems. The Stock-Yogo statistic addresses the problems of multiple included endogenous variables and distribution theory, but is likely a conservative test, as Stock and Yogo point out. Overall, we view the set of results as unable to reject the hypothesis of weak instruments, although the evidence is hardly compelling. In the next sub-section, we will employ ML-based techniques to develop instruments that we know a priori to be strong, in order to more conclusively determine if weak instruments account for the inability of GMM to center on the true parameter values in finite samples.

4.3. Where to Find Strong Instruments

Recall that one may express a maximum likelihood estimator as an equivalent instrumental variables estimator, for the appropriate choice of instruments. Here, our maximum likelihood estimator is implicitly using "optimal" instruments for expected output and the real interest rate. That is, the solution technique that is imbedded in the likelihood estimator computes the solutions for expected output and the real rate that are consistent with the posited model structure (and auxiliary VAR equations for inflation and the funds rate). One may write the optimal instruments in this case as

$$Z_t^o = B\underline{Y}_{t-1}. \tag{7}$$

where B is the matrix of reduced-form solution coefficients that constitute the unique, stable solution to the rational expectations model and $\underline{Y}_{t-1}$ is the matrix of lags of all the variables in the model cast in first-order form (see Anderson and Moore, 1985, and the appendix).[16] GMM, in contrast, simply forms linear projections of these variables on the instrument set. As the results in Table 5 suggest, the potential instrument projections available in this specification, given

---

[16]We include auxiliary equations in the model that define $E_t y_{t+1}$ and $i_t - E_t \pi_{t+1}$, and then select the corresponding rows of $B\underline{Y}_{t-1}$ to form the optimal instruments for these variables.

the inclusion of the lagged output gap among the regressors, makes the remaining instruments too weak to produce unbiased estimates of both μ and β. In imposing the full set of dynamic constraints that are consistent with the rational expectations solution to the model, our ML estimator evidently provides sufficiently strong instruments to properly center the distribution of estimates on the true values.

As confirmation of this hypothesis, the final three rows in each panel of Table 4 present the results from another Monte Carlo exercise. In this exercise, the true model is as described in the Monte Carlo exercises above. In each case, we use the solution for the model to form "optimal" instruments in the sense defined above, and compute instrumental variable estimates using these optimal instruments. As the results indicate, the bias that was present using the conventional GMM methodology disappears completely. The distribution of estimates is properly centered on the true value, even for sample size 125 (approximately the size of the data sample that we employ above). Note that the efficiency of the estimates remains low relative to ML, as expected. But here, one can improve the distribution of estimates in the Monte Carlo exercise by imposing the uniqueness and stability constraints required for a rational expectations solution. Once this is done, the dispersion of GMM estimates is almost ten times smaller than the distribution that does not impose these constraints.

Thus, at least for these data sets and this model, GMM fails because it provides weak instruments for the included endogenous variables y_t and $i_t - E_t\pi_{t+1}$. This does not mean that it is impossible to develop unbiased estimates for the parameters of interest, as the ML results suggest throughout. Simple instrumental variables estimates that use the instruments from ML, which embody the rational expectations constraints implied by the model, properly center on the true value in finite samples of the size in our study.

5. Conclusion

The output Euler equation has become the mainstay of much macroeconomic research, and especially of monetary policy research, over the past five years. The degree to which agents appear to look forward in this specification is important because it has an important influence on the dynamic responses of output to changes in the economic environment. In particular, the speed and contour of the response to monetary policy actions can differ substantially depending on the assumed degree of forward-looking behavior. Of course, such differences can, in principle,

allow a researcher to distinguish between more and less forward-looking variants of the model.

This paper attempts to make just such a distinction. Overall, the findings of this paper support a relatively modest weight on expected output, and a sizable and empirically robust weight on lagged output. A corresponding conclusion finds that specifications with larger weight on lagged output develop more significant estimates of the real rate effect.

These conclusions depend heavily on the estimates obtained from maximum likelihood. We focus on these estimates because another key finding in this paper is that GMM estimates of the degree of forward-looking behavior in this specification are biased in finite samples for this specification. We explore the source of this bias, and find that it appears to lie in the inability of GMM to provide sufficiently strong instruments. We show further that maximum likelihood, by explicitly imposing all of the constraints implied by the model with rational expectations, implicitly provides instruments that are sufficiently strong to properly center the parameter distributions on the true values.

An alternative conclusion is that the popular but highly stylized New Keynesian specification for aggregate output is simply so misspecified that it is difficult to draw any conclusions about macro dynamics (see Mavroeidis 2003). In that regard, we wish to re-emphasize that our conclusions about the extent of forward- and backward-looking behavior are restricted to the simple but widely-used specification explored here.

A. Appendix on Methodology

A.1. Maximum Likelihood Estimation

This procedure maximizes the concentrated log-likelihood function for equation (2), coupled with reduced-form VAR equations for the federal funds rate and inflation (also see Estrella and Fuhrer, 2003). The model that comprises equation (2) and the auxiliary VAR equations falls in the class of linear rational expectations models that may be represented in the format

$$\sum_{i=-\tau}^{\theta} H_i x_{t+i} + \sum_{i=1}^{\theta} H_i E_t(x_{t+i}) = \epsilon_t, \tag{A.1}$$

where τ and θ are positive integers, x_t is a vector of variables, and the H_i are conformable n-square coefficient matrices, where n is the number of endogenous variables in the model. The coefficient matrices H_i are completely determined by a set of underlying structural parameters Θ. The expectation operator $E_t(\cdot)$ denotes mathematical expectation conditioned on the process history through period t. The random shock ϵ_t is independently and identically distributed $N(0, \Omega)$. Note that the covariance matrix Ω is singular whenever equation (A.1) includes identities. We use the AIM procedure of Anderson and Moore (1985) to solve out for expectations of the future in terms of expectations of the present and the past. These expectations are then substituted into equation (A.1) to derive a representation of the model that we call the observable structure,

$$\sum_{i=-\tau}^{\theta} S_i x_{t+i} = \epsilon_t. \tag{A.2}$$

Having obtained the observable structure, it is relatively straightforward to compute the value of the likelihood function given the data and parameter values. The likelihood is defined as

$$\mathcal{L} = T(\log|\mathcal{J}| - .5\log|\hat{\Omega}|), \tag{A.3}$$

where T is the sample size, $\mathcal{J}$ is the Jacobian of transformation (which is time-invariant by assumption), and Ω is the variance-covariance matrix of the structural residuals ϵ_t. Details on the computation of the Jacobian may be found in Estrella and Fuhrer (2003). The residuals for each time period $t = 1, \ldots, T$ are computed, and the residual covariance matrix is then computed as

$$\Omega = (1/T)\epsilon\epsilon'. \tag{A.4}$$

17

Maximum likelihood estimation consists of finding the parameter values Θ, implicit in the coefficient matrices H_i of equation (A.1), that maximize equation (A.3).

A.2. GMM Estimation

Equation (2) is linear in variables and parameters, and hence (the two-step) GMM estimation is straightforward. Let Z_t be a $T \times q$ matrix of eligible instrumental variables, with each of the q instruments assumed orthogonal to η_t. Define Y to be the $T \times 1$ vector containing the time series values of the output gap y, and X to be the $T \times 4$ matrix containing the time series values for $[y_{t-1}, y_{t-2}, y_{t+1}, r_t]$, where r_t refers to any of the alternative definitions of the real interest rate, the benchmark being $i_t - \pi_{t+1}$. Defining the parameter vector $\gamma = [\alpha_1, \alpha_2, \mu, \beta]$, the GMM estimator is

$$\hat{\gamma}_n = (X'_n Z \hat{W}_T Z' X_n)^{-1} X'_n Z \hat{W}_T Z' Y_n, \tag{A.5}$$

where $\hat{W}_T$ is a $q \times q$ consistent estimator of $[\sum_{j=-\infty}^{\infty} E(Z_t Z'_{t-j} \eta_t \eta'_{t-j})]^{-1}$. We use the Newey-West estimate of $\hat{W}_T$ with a lag length of 4.

A.3. Generating Monte Carlo Data

The model used to generate data for the Monte Carlo exercises comprises the simplified IS equation (4), along with the estimated reduced-form VAR equations for inflation and the federal funds rate that are used in the maximum likelihood estimation described in section 3.1. These equations constitute a linear dynamic rational expectations model that can be cast in the form of equation (A.1). We solve for the expectations as described above and solve for the model's observable structure as in equation (A.2). We assume that the structural shocks in equation (A.2) are normally distributed with mean zero and covariance matrix as estimated on the U.S. data in section 3.2. For each replication of data in the Monte Carlo, we draw a set of joint normal iid shocks from this distribution for the three equations. The first τ (lagged) observations for the output gap, inflation rate, and funds rate are initialized at their steady-state values, the shocks are fed into equation (A.2), and the model is dynamically simulated. We discard the first fifty observations from each replication to minimize the effect of initial conditions.

References

[1] Anderson, G. S., Moore, G., 1985. A linear algebraic procedure for solving linear perfect foresight models. Economic Letters 17, 247-52.

[2] Casares, M., McCallum, B., 2000. An optimizing IS-LM framework with endogenous investment. NBER Working Paper No. 7908.

[3] Christiano, L., Fitzgerald, T., 2003. The band pass filter. International Economic Review, forthcoming.

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

[5] Cogley, T., Nason, J., 1995. Output dynamics in real-business-cycle models. American Economic Review 85, 492-511.

[6] Dennis, R., 2003. New Keynesian Optimal-Policy Models: An Empirical Assessment. manuscript, Federal Reserve Bank of San Francisco.

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

[8] Estrella, A., Fuhrer, J.C., 2003. Monetary policy shifts and the stability of monetary policy models. Review of Economics and Statistics 85, 94-104.

[9] Fuhrer, J.C., 1997. The (un)importance of forward-looking behavior in price specifications. Journal of Money, Credit, and Banking, August, 338-350.

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

[11] Fuhrer, J.C., Moore, G.R., 1995. Monetary policy trade-offs and the correlation between nominal interest rates and real output. American Economic Review 85, 219-239.

[12] Fuhrer, J.C., Moore, G.R., Schuh, S., 1995. Estimating the linear-quadratic inventory model: maximum likelihood versus generalized method of moments. Journal of Monetary Economics 35, 115-158.

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

[14] Goodfriend, M., King, R.G., 1997. The new neoclassical synthesis and the role of monetary policy. NBER Macroeconomics Annual, 231-283.

[15] Hall, A.R., Inoue A., 2003. The large sample behaviour of the generalized method of moments estimator in misspecified models. Journal of Econometrics 114, 361-394.

[16] Hall, A.R., Rudebusch, G., Wilcox, D.W., 1996. Judging instrument relevance in instrumental variables estimation. International Economic Review 37, 283-298.

[17] Jondeau, E., Bihan, H.L., 2003. ML vs. GMM estimates of hybrid macroeconomic models. Manuscript, Bank of France.

[18] Koenig, E., 1993. Rethinking the IS in IS-LM: adapting Keynesian tools to non-Keynesian economies. Economic Review, Q3. Federal Reserve Bank of Dallas, 33-49.

[19] Mankiw, G., Reis, R., 2001. Sticky information: a model of monetary nonneutrality and structural slumps. NBER working paper no. 8614.

[20] Mavroeidis, Sophocles, 2003. Identification and misspecification issues in forward-looking models. Manuscript, University of Amsterdam.

[21] McCallum, B.T., Nelson, E., 1999. Nominal income targeting in an open-economy optimizing model. Journal of Monetary Economics 43, 553-578.

[22] Nelson, C., Startz, R., 1990a. Some further results on the exact small sample properties of the instrumental variables estimator. Econometrica 58, 967-976.

[23] Nelson, C., Startz, R., 1990b. The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. Journal of Business 63, 5125-5140.

[24] Oliner, S., Rudebusch, G.D., Sichel, D., 1995. New and old models of business investment: a comparison of forecasting performance. Journal of Money, Credit, and Banking 27, 806-826.

[25] Oliner, S., Rudebusch, G.D., Sichel, D., 1996. The Lucas Critique revisited: assessing the stability of empirical Euler equations for investment. Journal of Econometrics 70, 291-316.

[26] Rotemberg, J., Woodford, M., 1999. In: Taylor, J.B. (Ed.), Interest rate rules in an estimated sticky price model. Monetary Policy Rules. University of Chicago Press, Chicago, IL, 57-119.

[27] Rudd, J., Whelan, K., 2003. Can rational expectations sticky-price models explain inflation dynamics? Manuscript, Federal Reserve Board.

[28] Rudebusch, G.D., 2002. Assessing nominal income rules for monetary policy with model and data uncertainty. Economic Journal 112, 1-31.

[29] Shea, J., 1997. Instrument relevance in multivariate linear models: a simple measure. Review of Economics and Statistics 79, 348-352.

[30] Smets, F., 2000. What horizon for price stability? European Central Bank Working paper no. 24.

[31] Staiger, D., Stock, J.H., 1997. Instrumental variables regression with weak instruments. Econometrica 65, 557-586.

[32] Stock, J.H., Wright, J., Yogo, M., 2002. A survey of weak instruments and weak identification in generalized method of moments. Journal of Business and Economic Statistics 20, 518-529.

[33] Stock, J.H., Yogo, M., 2003. Testing for weak instruments in linear IV regression. Unpublished manuscript, Harvard University.

[34] Svensson, L.E.O., 1999. Inflation targeting: some extensions. Scandinavian Journal of Economics 101, 337-361.

[35] Walsh, C.E., 1998. Monetary Theory and Policy. MIT Press, Cambridge, MA.

[36] West, K.D., Wilcox, D.W., 1994. Some evidence on the finite sample behavior of an instrumental variables estimator of the linear quadratic inventory model. In: Fiorito, R. (Ed.), Inventory Cycles and Monetary Policy. Berlin: Springer-Verlag, 253-282.

[37] Woodford, M., 1996. Control of the public debt: a requirement for price stability? NBER Working Paper No. 5984.

Table 1
ML Estimates of Output Euler Equation

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \mu E_{t-\tau} y_{t+1} - \beta E_{t-\tau}\left[\frac{1}{\kappa}\Sigma_{j=0}^{\kappa-1}\left(i_{t+j+m} - \pi_{t+j+m+1}\right)\right] + \eta_t$$

Model specification				Coefficient estimates and standard errors					LR test p-value	
Trend	τ	κ	m	$\alpha_1 + \alpha_2$	μ	SE(μ)	β	SE(β)	$\mu = 0$	$\beta = 0$
HP	0	1	0	0.53	0.44	0.04	0.002	0.004	0.67	0.62
HP	0	4	0	0.73	0.01	0.05	0.016	0.027	0.62	0.98
HP	0	40	0	0.74	0.00	0.05	0.038	0.090	0.98	0.59
HP	1	1	0	0.53	0.44	0.04	0.002	0.004	0.67	0.62
HP	1	4	0	0.61	0.27	0.04	0.008	0.013	0.63	0.91
HP	1	40	0	0.74	0.00	0.05	0.038	0.090	0.91	0.58
HP	0	1	-1	0.75	0.00	0.05	0.024	0.023	1.00	0.37
HP	0	4	-1	0.74	0.00	0.05	0.019	0.026	0.37	1.00
HP	0	40	-1	0.75	0.00	0.05	0.081	0.084	1.00	0.51
BP	0	1	0	0.63	0.37	0.00	0.001	0.001	0.03	0.01
BP	0	4	0	0.63	0.37	0.00	0.001	0.001	0.01	0.06
BP	0	40	0	0.63	0.37	0.01	0.005	0.002	0.06	0.01
BP	0	1	-1	0.91	0.00	0.03	0.013	0.006	1.00	0.08
BP	0	4	-1	0.64	0.36	0.01	0.001	0.000	0.08	0.10
BP	0	40	-1	0.89	0.00	0.03	0.097	0.030	0.10	0.01
Seg.	0	1	0	0.56	0.44	0.02	0.007	0.005	0.08	0.00
Seg.	0	4	0	0.57	0.43	0.03	0.010	0.006	0.00	0.20
Seg.	0	40	0	0.63	0.38	0.03	0.140	0.054	0.20	0.00
Seg.	0	1	-1	0.94	0.00	0.04	0.108	0.027	1.00	0.00
Seg.	0	4	-1	0.94	0.00	0.04	0.114	0.029	1.00	0.00
Seg.	0	40	-1	0.99	0.00	0.04	0.690	0.230	1.00	0.00
Quad.	0	1	0	0.55	0.45	0.02	0.005	0.003	0.00	0.00
Quad.	0	4	0	0.55	0.45	0.02	0.006	0.004	0.00	0.01
Quad.	0	40	0	0.58	0.41	0.03	0.071	0.031	0.01	0.00
Quad.	0	1	-1	0.92	0.00	0.04	0.101	0.027	1.00	0.00
Quad.	0	4	-1	0.55	0.45	0.02	0.005	0.003	0.00	0.01
Quad.	0	40	-1	0.91	0.00	0.04	0.554	0.133	1.00	0.00
CBO	0	1	0	0.54	0.46	0.01	0.005	0.002	0.00	0.00
CBO	0	4	0	0.54	0.46	0.02	0.006	0.003	0.00	0.00
CBO	0	40	0	0.56	0.44	0.03	0.050	0.030	0.00	0.00
CBO	0	1	-1	0.91	0.00	0.04	0.091	0.025	1.00	0.00
CBO	0	4	-1	0.54	0.46	0.02	0.005	0.002	0.00	0.00
CBO	0	40	-1	0.89	0.00	0.04	0.590	0.146	1.00	0.00

Note: The model specification columns provide the output trend procedure, the timing of expectations, and the timing and duration of the real rate. The hypothesis test columns report likelihood-ratio p-values for the given null hypothesis.

Table 2
GMM Estimates of Output Euler Equation

$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \mu E_{t-\tau} y_{t+1} - \beta E_{t-\tau}[\tfrac{1}{\kappa}\Sigma_{j=0}^{\kappa-1} (i_{t+j+m} - \pi_{t+j+m+1})] + \eta_t$$

Model specification				Coefficient estimates and standard errors					J-test
Trend	Instruments	κ	m	$\alpha_1 + \alpha_2$	μ	SE(μ)	β	SE(β)	p-value
HP	y, i, π (1 - 4)	1	0	0.45	0.65	0.15	-0.003	0.025	0.66
HP	y, i, π (1 - 4)	4	0	0.45	0.65	0.15	-0.006	0.024	0.68
HP	y, i, π (1 - 4)	1	-1	0.44	0.67	0.15	-0.006	0.018	0.67
HP	y, i, π (1 - 4)	4	-1	0.44	0.66	0.15	-0.005	0.025	0.67
HP	exog. (1 - 4)	1	0	0.49	0.56	0.08	-0.017	0.024	0.45
HP	exog. (1 - 4)	4	0	0.46	0.63	0.10	-0.018	0.027	0.37
HP	exog. (1 - 4)	1	-1	0.49	0.57	0.08	-0.023	0.023	0.51
HP	exog. (1 - 4)	4	-1	0.48	0.58	0.09	-0.019	0.025	0.46
HP	exog. (0 - 4)	1	0	0.45	0.60	0.09	0.022	0.025	0.27
HP	exog. (0 - 4)	4	0	0.46	0.61	0.10	0.010	0.027	0.25
HP	exog. (0 - 4)	1	-1	0.45	0.62	0.09	0.018	0.023	0.26
HP	exog. (0 - 4)	4	-1	0.47	0.58	0.09	0.016	0.025	0.25
BP	y, i, π (1 - 4)	1	0	0.56	0.46	0.02	0.001	0.003	0.01
BP	y, i, π (1 - 4)	4	0	0.55	0.47	0.02	0.000	0.003	0.01
BP	exog. (1 - 4)	1	0	0.64	0.36	0.03	0.008	0.004	0.78
BP	exog. (1 - 4)	4	0	0.64	0.36	0.04	0.007	0.004	0.77
Seg.	y, i, π (1 - 4)	1	0	0.52	0.50	0.05	0.010	0.013	0.40
Seg.	y, i, π (1 - 4)	4	0	0.53	0.48	0.05	0.008	0.016	0.48
Seg.	exog. (0 - 4)	1	0	0.51	0.52	0.05	0.008	0.018	0.34
Seg.	exog. (0 - 4)	4	0	0.48	0.55	0.05	0.014	0.018	0.28
Quad.	y, i, π (1 - 4)	1	0	0.51	0.51	0.06	0.016	0.016	0.35
Quad.	y, i, π (1 - 4)	4	0	0.53	0.49	0.05	0.013	0.020	0.43
Quad.	exog. (0 - 4)	1	0	0.48	0.55	0.07	0.016	0.020	0.34
Quad.	exog. (0 - 4)	4	0	0.44	0.59	0.07	0.024	0.020	0.32
CBO	y, i, π (1 - 4)	1	0	0.51	0.51	0.06	0.013	0.014	0.38
CBO	y, i, π (1 - 4)	4	0	0.52	0.49	0.05	0.010	0.017	0.46
CBO	exog. (0 - 4)	1	0	0.49	0.54	0.06	0.013	0.019	0.36
CBO	exog. (0 - 4)	4	0	0.46	0.58	0.06	0.019	0.019	0.31

Note: The model specification columns provide the output trend procedure, the timing and duration of the real rate, and the instrument set—both the variables used and, in parentheses, the number of lags.

Table 3
Tests for Normality of Estimated Maximum Likelihood Residuals

(CBO output gap, baseline specification, real rate $= i_t - E_t \pi_{t+1}$)

Test	p-value
Output Gap, 1966:Q1-2000:Q4	
Lilliefors (D)	> 0.1
Cramer-von Mises (W2)	0.2909
Watson (U2)	0.2827
Anderson-Darling (A2)	0.2073
Inflation, 1966:Q1-2000:Q4	
Lilliefors (D)	> 0.1
Cramer-von Mises (W2)	0.3800
Watson (U2)	0.3407
Anderson-Darling (A2)	0.3210
Funds rate, 1966:Q1-1979:Q3	
Lilliefors (D)	> 0.1
Cramer-von Mises (W2)	0.2894
Watson (U2)	0.2572
Anderson-Darling (A2)	0.1436
Funds rate, 1983:Q1-2000:Q4	
Lilliefors (D)	> 0.1
Cramer-von Mises (W2)	0.1888
Watson (U2)	0.2315
Anderson-Darling (A2)	0.1435

Table 4
Properties of ML and GMM estimators

Panel A. Estimates of expectational parameter μ

Estimation Method	Sample size	μ^T	Mean($\hat{\mu}$)	Median($\hat{\mu}$)	F(μ^T)	Median SE($\hat{\mu}$)
GMM	125	0.1	0.29	0.30	0.12	0.160
GMM	500	0.1	0.17	0.17	0.26	0.110
GMM	1000	0.1	0.14	0.14	0.31	0.080
GMM	5000	0.1	0.11	0.11	0.43	0.038
GMM	125	0.5	0.48	0.48	0.57	0.120
GMM	125	0.9	0.76	0.76	0.80	0.180
ML	125	0.1	0.13	0.11	0.48	0.089
ML	125	0.5	0.50	0.50	0.47	0.054
ML	125	0.9	0.90	0.90	0.48	0.050
Optimal Inst. GMM	125	0.1	0.07	0.12	0.47	0.250
Optimal Inst. GMM	125	0.5	0.51	0.50	0.48	0.140
Optimal Inst. GMM	125	0.9	0.91	0.89	0.51	0.220

Panel B. Estimates of interest rate sensitivity β

Estimation Method	Sample size	μ^T	Mean($\hat{\beta}$)	Median($\hat{\beta}$)	F(β^T)	Median SE($\hat{\beta}$)
GMM	125	0.1	0.30	0.3	0.84	0.180
GMM	125	0.5	0.49	0.49	0.52	0.180
GMM	125	0.9	0.50	0.48	0.55	0.200
GMM	500	0.1	0.44	0.43	0.73	0.120
GMM	1000	0.1	0.47	0.46	0.67	0.091
GMM	5000	0.1	0.50	0.49	0.57	0.043
ML	125	0.1	0.47	0.47	0.50	0.110
ML	125	0.5	0.52	0.51	0.46	0.098
ML	125	0.9	0.51	0.51	0.47	0.068
Optimal Inst. GMM	125	0.5	0.55	0.50	0.50	0.280
Optimal Inst. GMM	125	0.5	0.51	0.51	0.47	0.210
Optimal Inst. GMM	125	0.5	0.54	0.52	0.46	0.240

Note: The true data generating process has a $\mu = \mu^T$, which is displayed in the third column, and a $\beta = \beta^T = 0.5$ in all cases. F(x^T) is the empirical cdf of the $\hat{x}$ estimates evaluated at x^T.

Table 5
Instrument Relevance Tests

Model Specification		First-stage F-statistics		Partial R^2		Stock-Yogo
Trend	Instruments	F-stat(y)	F-stat(ρ)	$R^2(y)$	$R^2(\rho)$	$\min(\mathrm{eig}(G_T))$
HP	y, i, π (1 - 4)	7.0	36.8	0.10	0.88	1.4
HP	exog. (1 - 4)	5.0	8.0	0.19	0.79	1.6
BP	y, i, π (1 - 4)	140.7	31.6	0.55	0.88	15.0
BP	exog. (1 - 4)	83.2	7.8	0.57	0.80	8.3
Seg.	y, i, π (1 - 4)	100.0	37.2	0.26	0.78	1.9
Seg.	exog. (1 - 4)	56.1	14.9	0.28	0.82	1.7
Quad.	y, i, π (1 - 4)	77.4	34.6	0.29	0.79	2.7
Quad.	exog. (1 - 4)	46.0	9.1	0.34	0.78	2.2
CBO	y, i, π (1 - 4)	66.0	34.0	0.27	0.70	3.2
CBO	exog. (1 - 4)	40.2	8.0	0.34	0.70	2.4

Note: The model specification columns provide the definition of the output trend and the instrument set. Instruments are either a set of endogenous or exogenous variables with the number of lags included in parentheses.