Assessing Generalized Method-of-Moments Estimates of the Federal Reserve Reaction Function

Eric JONDEAU
Centre de recherche, Banque de France, and ERUDITE, Université Paris 12 Val-de-Marne, France (*eric.jondeau@banque-france.fr*)

Hervé LE BIHAN
Centre de recherche, Banque de France, Paris, France (*herve.lebihan@banque-france.fr*)

Clémentine GALLÈS
GREMAQ, Université des Sciences Sociales, Toulouse, France (*c_galles@hotmail.com*)

Estimating a forward-looking monetary policy rule by the generalized method of moments (GMM) has become a popular approach. We reexamine estimates of the Federal Reserve reaction function using several GMM estimators and a maximum likelihood (ML) estimator. First, we show that over the baseline period 1979–2000, these alternative approaches yield substantially different parameter estimates. Using Monte Carlo simulations, we show that the finite-sample GMM bias can account for only a small part of the discrepancy between estimates. We find that this discrepancy is more plausibly rationalized by the serial correlation of the policy shock, causing misspecification of GMM estimators through lack of instrument exogeneity. This correlation pattern is related to a shift in the reaction function parameters around 1987. Reestimating the reaction function over the 1987–2000 period produces GMM estimates that are very close to the ML estimate.

KEY WORDS: Finite-sample properties; Forward-looking model; Generalized method of moments; Maximum likelihood estimator; Monetary policy reaction function.

1. INTRODUCTION

According to the benchmark Taylor rule, central banks set the short-term interest rate in proportion of the inflation rate and the output gap. Since Taylor's (1993) prominent contribution, an abundant empirical as well as theoretical literature has claimed that central banks may have a forward-looking behavior (Clarida and Gertler 1997; Clarida, Galí, and Gertler 1998, 2000). This assumption involves using an adequate estimation method to overcome the presence of expected inflation in the policy rule. Following the influential work by Clarida et al. (1998), a large number of studies have thus used the generalized method of moments (GMM) to estimate forward-looking reaction functions (see, e.g., Mehra 1999; Clarida et al. 2000, henceforth CGG; Orphanides 2001; Rudebusch 2002).

In this article we reexamine the estimation of the Federal Reserve forward-looking policy rule and present some original empirical results. Although this topic has been widely studied, there are at least two motivations for our additional investigation. First, a large body of research over the last decade has analyzed the properties of GMM estimators and has produced numerous results that so far have not been incorporated in the estimation of policy rules. In particular, numerous authors have studied the small-sample properties of GMM estimators, in very different contexts (see the 1996 special issue of *JBES*, vol. 14, no. 3). These authors provided evidence that GMM estimators may be strongly biased and widely dispersed in small samples. Fuhrer, Moore, and Schuh (1995) also pointed out the poor finite-sample performance of GMM as compared with that of maximum likelihood (ML). In addition, several alternative GMM estimators have been proposed (Ferson and Foerster 1994; Hansen, Heaton, and Yaron 1996) that have very different small-sample properties. Moreover, alternative procedures for computing the GMM weighting matrix are likely to provide contrasting results. One of our aims is to reexamine estimates of the Federal Reserve reaction function in light of these developments. The forward-looking reaction function may be seen as an original field for investigating GMM properties.

A second strong motivation to investigate the Federal Reserve reaction function is the issue of parameter stability. The tenures of Paul Volcker (1979–1987) and Alan Greenspan (since 1987) as chairman of the Board of Governors of the Federal Reserve System have been characterized by two very contrasting subperiods in terms of interest rate movements, but no consensus has yet emerged on whether these two eras represent a single policy regime. Some authors have argued that there has been no significant difference in the way monetary policy is conducted since Volcker was appointed chairman in 1979. In particular, CGG found that during the Volcker–Greenspan period, the Federal Reserve has adopted a proactive stance toward controlling inflation. Other authors (Judd and Rudebusch 1998; Rudebusch 2002) concentrated on Greenspan's tenure. Estimations in this article are found to provide some insight into this issue.

The article is organized as follows. In Section 2 we review the specification of the forward-looking monetary policy reaction function. In Section 3 we provide evidence that alternative

Journal of Business & Economic Statistics
April 2004, Vol. 22, No. 2
DOI 10.1198/073500104000000118

GMM and ML estimation procedures produce very contrasting estimates of the reaction function over the 1979–2000 period; however, a standard diagnostic check based on both GMM and ML estimates does not point to a model misspecification problem. In Section 4 we show that the dispersion in reaction function parameters cannot be imputed to the previously reported poor finite-sample properties of GMM estimators. In Section 5 we present evidence that the reaction function parameters are likely to have shifted following the change in the Federal Reserve chairman in 1987:Q3. We suggest that this shift is responsible for the large discrepancy between estimation procedures and also that a shift in variance explains why standard specification tests did not detect this misspecification. Finally, in Section 6, we find that estimates over the 1987:Q3–2000 period are remarkably close one to each other. In particular, the inflation parameter is estimated at a much lower value than that obtained in some studies over a long sample. Details on the estimation procedures are relegated in an Appendix available from the authors.

2. THE FORWARD–LOOKING REACTION FUNCTION

According to the baseline policy rule proposed by Taylor (1993), the central bank is assumed to set the nominal short-term interest rate (i_t) as a function of the (four-quarter) inflation rate ($\overline{\pi}_t$) and the output gap measure (y_t). As a matter of fact, most empirical studies investigated "modified" Taylor rules. First, central banks appear to smooth changes in interest rates, a behavior found to be motivated on theoretical grounds (e.g., Sack and Wieland 2000). Second, several authors have claimed that the behavior of central banks is consistent with a forward-looking reaction function, in which the interest rate level is set as a function of expected inflation and output gap (Clarida and Gertler 1997; Clarida et al. 1998, 2000; Mehra 1999; Orphanides 2001). Consequently, we specify the empirical policy rule as a partial-adjustment model, in which the short-term rate adjusts gradually to its target, defined in terms of future annual inflation and output gap,

$$i_t = \rho(L)i_{t-1} + \big(1-\rho(1)\big)\big(i^* + \beta(E_{t-1}\overline{\pi}_{t+4} - \pi^*) + \gamma E_{t-1}y_{t+1}\big) + \eta_t, \quad (1)$$

where i^* is the long-run equilibrium nominal interest rate and π^* is the inflation target. (The output gap target is assumed to be 0.) The term η_t can be interpreted as a random policy shock, which may reflect the central bank's failure to keep the interest rate at the level prescribed by the rule or the deliberate decision to deviate transitory from the rule. Note that we assume this shock to be serially uncorrelated for the lags of the interest rate to be valid instruments in the GMM approach. We define $\rho(L) = \rho_1 + \rho_2 L + \cdots + \rho_m L^{m-1}$, with $\rho(1)$ measuring the degree of interest rate smoothing. Later we consider one or two lags, depending on the estimation period.

The expression E_{t-1} denotes the expectation operator conditional on the information set $I_{t-1} = \{i_{t-1}, \pi_{t-1}, y_{t-1}, \dots\}$. This dating of expectations is consistent with current inflation and output gap not being observed in real time. It may be argued that the current interest rate is actually included in the central bank's information set, because it is its control variable. However, for the econometrician, because the current interest rate is evidently correlated to the error term, it cannot be used as an instrument in the GMM approach. Therefore, for convenience, we include only lagged variables in the information set, an assumption that is not restrictive from an econometrician's stand point. Equation (1) is one of the baseline specifications estimated by CGG, Orphanides (2001), or Rudebusch (2002).

The long-run interest rate and inflation targets (i^* and π^*) are not identified, because the constant term is proportional to $\alpha = i^* - \beta\pi^* = r^* + (1-\beta)\pi^*$, with $r^* = i^* - \pi^*$ the equilibrium real rate. Although r^* and π^* are not separately identifiable in a single-equation approach, an estimate of the equilibrium real rate can be obtained from an auxiliary model, including, for instance, an I–S curve (see Sec. 3.3).

3. A 1979–2000 REACTION FUNCTION: EVIDENCE FROM ALTERNATIVE ESTIMATORS

3.1 Data

We analyze the Federal Reserve monetary policy reaction function over the period 1979:Q3–2000:Q3 using quarterly data drawn from the OECD databases BSDB and MEI. The Federal funds rate is used as the monetary policy instrument. Inflation is defined the rate of growth of the gross domestic product (GDP) deflator (denoted by P_t), so that $\pi_t = 400(\ln(P_t) - \ln(P_{t-1}))$ and $\bar{\pi}_t = \frac{1}{4}\sum_{i=0}^{3}\pi_{t-i}$. Output gap is defined by the percent deviation of real GDP (Q_t) from potential GDP (Q_t^*), that is, $y_t = 100(\ln(Q_t) - \ln(Q_t^*))$. Following a number of recent studies (CGG; Rudebusch 2002), we use the output gap series constructed by the Congressional Budget Office.

Our sample period covers the tenures of Volcker (1979:Q3–1987:Q2) and Greenspan (1987:Q3 to now). Using similar data, CGG, Mehra (1999), and Estrella and Fuhrer (2003) obtained statistical evidence that the reaction function is essentially stable over the two tenures. Other studies (e.g., Orphanides 2001; Rudebusch 2002) prefer to focus on Greenspan's tenure, raising the issue of parameter stability. In line with theoretical models of monetary policy rules, these empirical studies maintain the assumption that the interest rate and inflation are stationary processes. In a first stage, we assume that parameters are stable over the whole period and, although empirical evidence is not clear cut (see Fuhrer and Moore 1995), that the model is stationary. In Section 5 we address the issue of stability of the policy rule parameters over the two tenures. Therefore, we introduce the possibility of a shift in the long-run equilibrium nominal interest rate and in the target inflation rate in 1987:Q3, thus weakening the assumption of stationarity.

3.2 GMM Estimates

As is well known, estimating (1) by ordinary least squares (OLS) using actual future realizations in place of expectated terms would provide inconsistent estimators, because the associated error term $\varepsilon_t = \eta_t - (1-\rho(1))(\beta(\overline{\pi}_{t+4} - E_{t-1}\overline{\pi}_{t+4}) + \gamma(y_{t+1} - E_{t-1}y_{t+1}))$ would be correlated with endogenous regressors. This problem can be overcome by using the GMM. This technique requires only that the error term ε_t be orthogonal to a vector of instruments $\mathbf{Z}_{t-1}$ belonging to the information set I_{t-1}, so that $E[\varepsilon_t \mathbf{Z}_{t-1}] = E[\mathbf{g}_t(\boldsymbol{\theta})] = \mathbf{0}$, where

$\boldsymbol{\theta}$ denotes the vector of unknown parameters. An efficient GMM estimator of $\boldsymbol{\theta}$ is obtained by minimizing the expression $\bar{\mathbf{g}}(\boldsymbol{\theta})'(\mathbf{S}_T)^{-1}\bar{\mathbf{g}}(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$, where $\bar{\mathbf{g}}(\boldsymbol{\theta}) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{g}_t(\boldsymbol{\theta})$ and $\mathbf{S}_T$ is a consistent estimator of the long-run covariance matrix of $\mathbf{g}_t(\boldsymbol{\theta})$. Provided that instruments are correlated with endogenous regressors and uncorrelated with the error term, GMM estimators are strongly consistent and asymptotically normal (Hansen 1982). Early studies on the use of GMM in a rational-expectation framework were conducted by Cumby, Huizinga, and Obstfeld (1983), Hansen and Singleton (1982), and Hayashi and Sims (1983). In this context, the GMM approach is very appealing, because it requires only identifying relevant instruments and does not involve strong assumptions on the underlying model.

We consider three alternative GMM estimators proposed in the theoretical literature: the two-step GMM, the iterative GMM, and the continuous-updating GMM. To our knowledge, all estimations of the forward-looking reaction function based on GMM have so far relied on the two-step estimator. These approaches differ in the way in which the parameter vector and the long-run covariance matrix interact. Two-step GMM is the two-step, two-stage least squares procedure initially proposed by Hansen (1982) and Cumby et al. (1983). The iterative approach, suggested by Ferson and Foerster (1994) and Hansen et al. (1996), relies on estimating the parameter vector $\boldsymbol{\theta}$ and the covariance matrix $\mathbf{S}_T$ iteratively. In the continuous-updating GMM approach, developed by Hansen et al. (1996) and studied by Stock and Wright (2000), Newey and Smith (2003), and Ma (2002), the parameter vector and the covariance matrix are determined in the minimization simultaneously.

We consider four variants of the covariance matrix estimator. $\mathbf{S}_{1T}$ is the estimator proposed by Newey and West (1987) with bandwidth parameter $L = 4$ and a decreasing-weight Bartlett kernel. Because this estimator is likely to provide inconsistent estimates of $\mathbf{S}_T$, we also consider two estimators with data-dependent bandwidth: $\mathbf{S}_{2T}$, proposed by Andrews and Monahan (1992), and, $\mathbf{S}_{3T}$, proposed by Newey and West (1994). Whereas the former is based on an AR(1) process for the moment conditions, the bandwidth of the latter is computed nonparametrically. Finally, $\mathbf{S}_{4T}$ is the estimator suggested by West (1997) with $L = 4$ and a rectangular kernel. This estimator makes use of the fact that the error term has a moving average [MA(4)] structure. An Appendix available from the authors provides further details on the GMM procedures.

Before proceeding, we address the issue of instrument choice. Our instrument set includes four lagged values of the interest rate, inflation, and the output gap: $i_{t-1}, \overline{\pi}_{t-1}, y_{t-1}, \ldots, i_{t-4}, \overline{\pi}_{t-4}, y_{t-4}$. This set contrasts with most previous GMM estimates of the Federal Reserve reaction function, which include several additional macroeconomic variables as instruments. There are two motivations for choosing such a restricted information set. First, when a large number of instruments is selected, some instruments may be weakly relevant or redundant, thus deteriorating the finite-sample properties of GMM (see the discussion in Sec. 4). Second, the comparison of alternative GMM and ML estimators using Monte Carlo simulations (as performed in Sec. 4) necessitates a plausible data-generating process for all instruments. We therefore intentionally reduce the information set to lags of the three variables used in the structural model discussed later.

Table 1 reports parameter estimates of the reaction function. We first consider the two-step GMM with covariance matrix estimator $\mathbf{S}_{1T}$ (first rows of panel A) with only one lagged interest rate ($\rho_2 = 0$). This case corresponds to the estimation method adopted by CGG (Table 4, second row), and is considered throughout the article as the baseline estimate. The estimate of the response to expected inflation ($\beta = 2.63$) is significantly larger than the coefficient of 1.5 originally proposed by Taylor (1993). The estimate of the output gap parameter ($\gamma = .71$) is very close to Taylor's coefficient of .5, although it is only weakly significant. These estimates are also close to those obtained by CGG over the period 1979–1996. In addition, reestimating equation (1) using the same instrument set as CGG over our sample did not alter estimation results. Broadly speaking, the standard, two-step GMM approach provides point estimates that are rather robust to slight changes in the specification.

The iterative and continuous-updating GMM (panels B and C) yield even larger estimates for the inflation parameter. Estimate of β is as high as 3.59 for the iterative GMM and 3.62 for the continuous-updating GMM. The values are larger than most of those found in the empirical policy rule literature. The output gap parameter is lower than with two-step GMM and it is statistically insignificant.

We turn now to the choice of the long-run covariance matrix. The broad picture suggested by our results is that the different covariance matrices provide widely dispersed point estimates. Broadly speaking, estimates obtained using $\mathbf{S}_{1T}$ and $\mathbf{S}_{4T}$ (which impose the strongest constraints on the covariance matrix) are close to one another. Similarly, estimators based on $\mathbf{S}_{2T}$ and $\mathbf{S}_{3T}$ (which do not make any assumption on the order of the MA process) are very similar. But the two groups differ significantly. For instance, the inflation parameter estimated with continuous-updating GMM decreases from 3.62 with estimator $\mathbf{S}_{1T}$ or 3.48 with $\mathbf{S}_{4T}$ to 2.72 with $\mathbf{S}_{2T}$ and 2.11 with $\mathbf{S}_{3T}$.

The dispersion between estimation approaches also transpires in Hansen's J statistics. If one were to use the covariance matrices $\mathbf{S}_{2T}$ or $\mathbf{S}_{3T}$, one would strongly reject the overidentifying restrictions, on the basis of the asymptotic p values. This suggests that some instruments fail to satisfy the orthogonality conditions. On the contrary, using $\mathbf{S}_{1T}$ or $\mathbf{S}_{4T}$, one would not be able to reject overidentifying restrictions. It should be noted, however, that due to the small size of our sample, the asymptotic distribution of the test statistics may fail to approximate the finite-sample distribution accurately. We thus computed the finite-sample distribution using Monte Carlo simulation as described in Section 4. Then, based on finite-sample p values, over-identifying restrictions are no longer rejected. Indeed, the two largest J statistics obtained in Table 1 (corresponding to the two-step and iterative GMM with the covariance matrix $\mathbf{S}_{3T}$) have size-adjusted p values equal to 35% and 11%.

On the whole, although the conventional J statistics do not point to model misspecification, GMM estimates appear to be widely dispersed. As suggested by Stock and Wright (2000, p. 1090), this may reflect the presence of weak instruments or a specification problem.

Table 1. GMM Estimates, 1979:Q3–2000:Q3

	Panel A: Two-step GMM		*Panel B: Iterative GMM*		*Panel C: Continuous-updating GMM*	
	Estimate	*SE*	*Estimate*	*SE*	*Estimate*	*SE*
Covariance-matrix estimator $\mathbf{S}_{1T}$						
ρ_1	.831	.043	.782	.061	.766	.053
β	2.631	.486	3.591	.566	3.619	.385
γ	.712	.388	.489	.372	.435	.309
α	−.585	1.465	−2.836	2.479	−2.982	2.277
J statistic (statistic/asymptotic p value)	9.765	.370	6.371	.702	6.180	.722
(size-adjusted p value)		.813		.867		.846
Covariance-matrix estimator $\mathbf{S}_{2T}$						
ρ_1	.875	.031	.938	.030	.837	.014
β	3.025	.449	6.000	2.296	2.725	.141
γ	1.377	.261	1.725	1.101	1.279	.337
α	−1.249	1.188	−8.117	6.833	−.369	.313
J statistic (statistic/asymptotic p value)	23.393	.005	14.006	.122	13.152	.156
(size-adjusted p value)		.293		.395		.327
Covariance-matrix estimator $\mathbf{S}_{3T}$						
ρ_1	.848	.025	.932	.024	.804	.018
β	2.568	.248	6.979	2.302	2.107	.412
γ	.937	.126	2.312	1.036	1.566	.420
α	−.208	.811	−11.254	6.854	1.755	.802
J statistic (statistic/asymptotic p value)	28.798	.001	29.057	.001	13.330	.148
(size-adjusted p value)		.348		.110		.370
Covariance-matrix estimator $\mathbf{S}_{4T}$						
ρ_1	.811	.041	.831	.044	.799	.054
β	2.507	.442	2.685	.496	3.475	.382
γ	.364	.242	.385	.292	.300	.262
α	−.241	1.201	−.601	1.463	−2.579	2.212
J statistic (statistic/asymptotic p value)	13.914	.177	7.982	.631	7.739	.561
(size-adjusted p value)		.313		.680		.678

3.3 ML Estimates

We now focus on the alternative ML estimation procedure. This approach requires that an auxiliary model be estimated to forecast the expected variables appearing in the reaction function (here the inflation rate and the output gap). An appealing advantage of ML over GMM, in a forward-looking context, is that expectations obtained with ML estimation are fully model consistent. The ML approach is evidently more demanding, because the auxiliary model must be estimated. However, in the present case, the widely used Phillips curve/I–S curve framework provides a reliable benchmark model of the inflation output joint dynamics.

The Auxiliary Model. The auxiliary model is inspired by the model of Rudebusch and Svensson (1999), which embodies the main features of the standard macroeconomic paradigm. The key relationships of the model are

$$\pi_t = \alpha_{\pi 1}\pi_{t-1} + \alpha_{\pi 2}\pi_{t-2} + \alpha_{\pi 3}\pi_{t-3} + \alpha_{\pi 4}\pi_{t-4} + \alpha_y y_{t-1} + u_t \qquad (2)$$

and

$$y_t = \beta_{y1} y_{t-1} + \beta_{y2} y_{t-2} + \beta_r(\bar{\imath}_{t-1} - \bar{\pi}_{t-1} - \beta_0) + v_t, \qquad (3)$$

where $\bar{x}_t = \frac{1}{4}\sum_{i=0}^{3} x_{t-i}$ denotes the four-quarter moving average of x_t. The Phillips curve (2) relates quarterly inflation (π_t) to its own lags and to lagged output gap. To preclude any inflation/output gap trade-off in the long run, we impose that the four autoregressive parameters must sum to 1. The constant term is also set to 0, so that the steady-state value of output gap is 0. Using the likelihood ratio (LR) test, this joint restriction is not rejected in our ML estimation (with a p value of .57). The I–S curve (3) relates the output gap to its own lags and to the four-quarter moving average of the short real rate. Parameter β_0 may be interpreted as the equilibrium real rate, because it is the value of the real rate consistent with a steady-state output gap equal to 0.

The backward-looking nature of this model can be pointed to as a potential source of misspecification. However, such a model has proved to be a robust representation of the U.S. economy. Moreover, no compelling empirical forward-looking counterpart of the Rudebusch–Svensson model has yet emerged (see Estrella and Fuhrer 2003). We checked the robustness of our auxiliary model by estimating the hybrid model proposed by Rudebusch (2002); our estimate of the reaction function was essentially unaltered within this framework.

An important issue concerning the auxiliary model is the stability of parameters over time. To investigate this issue, we performed a LR test of parameter stability on each equation, allowing an unknown breakpoint, as suggested by Andrews (1993) and Andrews and Ploberger (1994). In addition to the asymptotic p values, we used size-adjusted p values computed using a bootstrap procedure that reproduces the pattern of residuals. Due to the heteroscedasticity pattern found in the I–S residuals, we used a block bootstrap approach to capture serial dependence in conditional variance. We considered different block lengths, but our results were found to be robust to the choice of the block length. To save space, results of these stability tests are not reported here, but are available on request. Using asymptotic p values as well as size-adjusted p values, we do not reject the stability of the auxiliary equations. This result

Table 2. ML Estimates, 1979:Q3–2000:Q3

Reaction function			Phillips curve			I–S curve		
Parameter	*Estimate*	*SE*	*Parameter*	*Estimate*	*SE*	*Parameter*	*Estimate*	*SE*
ρ_1	.713	.060	$\alpha_{\pi 1}$	.501	.120	β_{y1}	1.113	.111
β	1.879	.272	$\alpha_{\pi 2}$	.022	.150	β_{y2}	−.192	.101
γ	.020	.271	$\alpha_{\pi 3}$	.401	.152	β_r	−.089	.056
α	1.368	1.312	$\alpha_{\pi 4}$	.076		β_0	3.193	1.227
			α_y	.127	.043			
	Statistic	*p value*		*Statistic*	*p value*		*Statistic*	*p value*
Q(4)	5.364	.252	Q(4)	.869	.929	Q(4)	5.337	.254
Q(8)	15.973	.043	Q(8)	7.070	.529	Q(8)	8.909	.350
R(4)	36.567	.000	R(4)	4.539	.338	R(4)	*19.543*	*.001*
R(8)	47.681	.000	R(8)	*10.394*	*.238*	R(8)	*36.886*	*.000*
J–B	125.284	.000	J–B	1.405	.495	J–B	25.104	.000
see	1.016		see	.793		see	.695	
log-L	−295.214							

is in accordance with the evidence reported by Rudebusch and Svensson (1999) and indicates some robustness of the model with respect to the Lucas critique.

ML Parameter Estimates. The complete model is solved using the generalized saddlepath procedure developed by Anderson and Moore (1985) and described in an available Appendix. The Phillips curve, the I–S curve, and the reaction function are estimated simultaneously, with a free covariance matrix of innovations. Stationarity of the model requires that the following constraints on parameters hold: $\beta > 1$, $\alpha_y > 0$, and $\beta_r < 0$ (see, e.g., Taylor 1999 or CGG). These boundary conditions are readily satisfied on our data, because the three parameters are estimated to be significantly different from their boundary value. In addition, the two largest eigenvalues associated with the reduced form of the model have moduli of .95, suggesting that shocks are likely to have a persistent, yet stationary, effect on the system.

Parameter estimates of model (1)–(3) are reported in Table 2. Reported standard errors are corrected for heteroscedasticity of residuals. The effect of output gap on inflation is quite large ($\alpha_y = .127$), and the response of output gap to the real rate is $\beta_r = -.089$. These effects are slightly lower than those obtained by Rudebusch and Svensson (1999) over the 1961–1996 period, but they are correctly signed and significantly different from 0. Summary statistics reveals that for both equations residuals are not serially correlated, although the I–S curve displays some heteroscedasticity.

Turning to the reaction function, parameter estimates contrast sharply with those obtained with GMM. The inflation parameter is found to be much lower than all point estimates obtained by GMM ($\beta = 1.88$). The estimate of the output gap parameter γ also differs markedly from GMM estimates, because it is essentially 0 and nonsignificant. Finally, the smoothing parameter ρ_1 is lower than the GMM estimate (.71 vs. .83). Residual analysis provides two main results. First, the Ljung–Box statistic, $Q(4)$, does not reject the null of no serial correlation up to lag 4, whereas $Q(8)$ rejects the null only marginally. Second, the Engle statistics for conditional heteroscedasticity [denoted by R(k) and computed from a regression of squared residuals on k own lags], indicates that residuals are strongly heteroscedastic. Interestingly, heteroscedasticity has already been pointed out by Sims (2001) to be a major feature of the Federal Reserve policy rule and as a possible source of misspecification of the reaction function. Finally, the Jarque–Bera test statistic, denoted by J–B, strongly rejects the normality of reaction function residuals.

4. INVESTIGATING FINITE–SAMPLE PROPERTIES

Here we investigate possible explanations for the wide discrepancy between estimates of the reaction function parameters. As pointed out by Hall and Rossana (1991) in a related context, two factors are likely to explain this discrepancy. First, both GMM and ML estimators may suffer from finite-sample biases. The GMM finite-sample bias has been analyzed by many authors (e.g., Nelson and Startz 1990; Hall, Rudebusch, and Wilcox 1996; Staiger and Stock 1997). This bias typically originates in weak instrument relevance (see, e.g., the survey by Stock, Wright, and Yogo 2002) or in instrument redundance (see Breusch, Qian, Schmidt, and Wyhowki 1999; Hall and Peixe 2003 for a discussion of this notion). Also, a downward bias on the autoregressive parameter occurs in partial-adjustment models, even when the model is correctly specified (Sawa 1978) or when estimators are designed so as to be immune to residual autocorrelation (Hall and Rossana 1991). Second, the model may be misspecified. A possible source of misspecification in our setup is the instability of the reaction function, which would yield an inconsistency of GMM and ML estimators. We first assess the finite-sample properties of GMM and ML estimators, using Monte Carlo simulations.

4.1 Experiment Design

Simulations are performed using the complete model (1)–(3), rewritten in the autoregressive companion form. Parameters and the covariance matrix are those stemming from ML estimation reported in Section 3.3. For a given sample size T, a sequence of $T + 50$ random innovations is drawn from an iid Gaussian distribution with mean 0 and the covariance matrix estimated over the sample. We set $T = 85$, corresponding to the actual sample size. Initial conditions are set equal to the sample averages, whereas the first 50 entries are discarded to reduce the effect of initial conditions on the solution path. The Monte Carlo experiment is based on $N = 2{,}000$ replications. For each artificial database, estimation is performed as follows. For GMM, the reaction function is estimated with four lags of (simulated) inflation, output gap, and interest rate as instruments. For ML, the complete model is estimated simultaneously. Two-step and iterative GMM estimators are obtained by simple matrix computations, whereas continuously updating GMM and ML estimators are obtained using a numerical optimization routine. (Simulations are performed with GAUSS version 3.2 on a Pentium III platform, using the BFGS algorithm of the CO procedure for constrained optimization. No discrepancies were found when using alternative algorithms. All estimations are performed using numerical derivatives.)

In some experiments, the continuous-updating GMM estimator failed to converge. Hansen et al. (1996) and Smith (1999) also reported an important number of crashes and some difficulties in obtaining reasonable parameter estimates with this estimator. Hence, in Table 3 two rows are devoted to the continuous-updating GMM. The first row reports distribution statistics after we discarded only estimates that reached the maximum number of iterations (here 200). The second row presents estimates satisfying the additional criterion that the smoothing parameter ρ_1 lies inside the interval $[-1, 1]$. For instance, in our Monte Carlo experiment with $\mathbf{S}_{1T}$, 8.3% of estimations failed to converge, whereas 3.4% fell outside of this parameter space, so that 11.7% of estimations were finally discarded. Note that we used the estimates from two-step GMM as starting values for the continuous-updating GMM; using random starting values would worsen the properties of this estimator.

4.2 Estimator Biases

The distributions of the GMM and ML estimators are summarized in Table 3. Because parameter α does not provide incremental insight into the finite-sample properties, we do not report results for this parameter. The dispersion of the distribution is measured with both standard deviation and median absolute deviation (MAD), to avoid misinterpretation in cases where the distribution is likely to have unbounded moments. Figures 1 and 2 display the distribution of parameters ρ_1, β, and γ for the ML and the two-step GMM (with $\mathbf{S}_{1T}$) procedures. The table and the figures reveal three main results regarding the parameter finite-sample properties. First, although the sample size is rather small, we obtain statistically significant, yet economically unimportant, biases. The autoregressive parameter is found to be slightly biased toward 0, consistent with the analytical result of Sawa (1978). For β, the bias is about .04 for GMM estimators, but as low as $-.017$ with ML. In addition, parameter γ is underestimated regarding of the estimation approach. One possible reason why GMM biases are found to be small relative to those obtained in other contexts (e.g., Fuhrer et al. 1995) is that weak instrument identification is not a key issue in the present context. The number of instruments is here quite small, and only two instruments (y_{t-3} and y_{t-4}) turn out to be asymptotically uncorrelated with endogenous variables, conditional on other variables in the instrument set.

Second, the dispersion of parameter estimates is much lower with ML than with GMM. For instance, for β, the MAD is .13 with ML and .21 with two-step GMM, .23 with iterative GMM, and .30 with continuous-updating GMM (with covariance matrix $\mathbf{S}_{1T}$). When "unreasonable" outcomes are excluded, the MAD of the truncated continuous-updating estimator is still as high as .27. The lower dispersion of ML compared with GMM reflects the incorporation of relevant cross-equation restrictions in the ML procedure.

Third, the distribution of continuous-updating GMM estimator is markedly fat-tailed and asymmetric. The autoregressive parameter is left-skewed, whereas the inflation parameter is right-skewed. Interestingly, the use of covariance matrix $\mathbf{S}_{4T}$ allows to dramatically reduce the dispersion of the continuous-updating estimator. With such a covariance matrix, it is barely more dispersed than other GMM estimators.

The finding that the continuous-updating GMM estimator has fat tails and yields a nonnegligible proportion of implausible estimates was previously reported by Hansen et al. (1996). The poor performance of the continuous-updating GMM was also documented by Smith (1999) and Stock and Wright (2000) for some of their estimates and Monte Carlo experiments. The overall characteristics of the continuous-updating GMM can be explained by two factors. (We are grateful to an anonymous referee for suggesting these explanations.) First, unlike two-step and iterative GMM, the objective function is not quasi-quadratic, because the parameter vector and the covariance matrix are determined simultaneously. Therefore, in a nonnegligible number of cases, the objective function is minimized at arbitrarily large values of the parameter vector (see the findings and discussions in Hansen et al. 1996; Stock and Wright 2000). Second, the numerical search algorithm for the minimizer fails sometimes. Because the objective function is not necessarily well behaved, quasi-Newton optimization methods, such as the BFGS algorithm, may converge to a local minimum or even diverge to infinity.

In our context, the poor performance of the continuous-updating estimator may be arguably exacerbated by two elements. First, the finite-sample properties may be worsened by the very small size of the sample. We found that in a similar Monte Carlo experiment with a sample size of 200, the continuous-updating estimator displays the same distributional properties as the other GMM estimators. Second, continuous-updating estimation is particularly unsuccessful with inconsistent ($\mathbf{S}_{1T}$) or nonparametric ($\mathbf{S}_{3T}$) covariance matrices. As suggested earlier, imposing relevant constraints on the covariance matrix (as with $\mathbf{S}_{4T}$) allows to considerably reduce the frequency of crashes (.9%) and unreasonable outcomes (.7%). This issue has been previously outlined by Burnside and Eichenbaum (1996), who argued in a similar context that the more constrained the covariance matrix, the less biased the Wald test statistics. On the whole, we may conclude that for small sample sizes, the difficulty in solving the optimization problem is not compensated for by the benefits of estimating the parameters and the covariance matrix simultaneously.

4.3 The J Statistic

The rightmost columns of Table 3 report the rejection rates and size-adjusted critical values associated with the J statistic. Rejection rates are the percentages of the 2,000 replications in

Table 3. Monte Carlo Simulation Asssuming That the Model Is Correctly Specified

	ρ_1 (true value: .71)				β (true value: 1.88)				γ (true value: .02)				*J statistic* Rejection rates			*J statistic* Size-adjusted critical values		
	Mean	*Std*	*Median*	*MAD*	*Mean*	*Std*	*Median*	*MAD*	*Mean*	*Std*	*Median*	*MAD*	*1%*	*5%*	*10%*	*1%*	*5%*	*10%*
GMM																		
Covariance-matrix estimator $\mathbf{S}_{1T}$																		
Two-step	.67	.11	.69	.07	1.96	.42	1.92	.21	−.01	.39	−.03	.22	.29	.47	.57	45.8	34.7	30.0
Iterative	.65	.13	.67	.08	1.95	.46	1.91	.23	−.03	.43	−.05	.23	.09	.21	.31	32.0	24.0	20.8
Continuous-updating (8.3%)	.52	.26	.60	.14	12.71	526.9	1.94	.30	−9.86	561.7	−.16	.29	.01	.06	.12	20.8	17.3	15.2
Truncated CU (11.7%)	.57	.20	.62	.12	1.96	2.13	1.93	.27	−.19	1.29	−.14	.27	.00	.06	.11	20.5	17.1	15.0
Covariance-matrix estimator $\mathbf{S}_{2T}$																		
Two-step	.67	.12	.69	.07	1.96	.43	1.92	.22	−.01	.42	−.03	.23	.35	.53	.64	54.4	39.3	32.4
Iterative	.63	.17	.66	.09	1.94	.48	1.91	.24	−.04	.44	−.07	.23	.11	.25	.36	34.4	25.9	22.6
Continuous-updating (1.8%)	.59	.19	.63	.11	4.06	89.88	1.93	.29	−5.59	103.2	−.12	.29	.04	.14	.24	25.5	20.9	18.2
Truncated CU (2.8%)	.60	.18	.63	.11	1.98	1.03	1.93	.28	−.14	.91	−.12	.28	.04	.14	.24	25.0	20.9	18.2
Covariance-matrix estimator $\mathbf{S}_{3T}$																		
Two-step	.67	.13	.69	.07	1.96	.49	1.92	.22	−.01	.48	−.04	.24	.52	.67	.74	176.6	82.4	56.9
Iterative	.62	.18	.66	.09	1.91	.58	1.90	.25	−.05	.77	−.07	.24	.22	.36	.45	89.5	41.2	30.2
Continuous-updating (12.9%)	.59	.22	.62	.12	25.75	1,292.0	1.92	.33	−41.8	1,175.2	−.13	.32	.08	.19	.30	37.7	23.9	20.2
Truncated CU (14.3%)	.59	.18	.63	.11	2.05	4.16	1.93	.31	−.13	2.36	−.13	.30	.05	.17	.28	28.4	21.9	19.2
Covariance-matrix estimator $\mathbf{S}_{4T}$																		
Two-step	.68	.11	.69	.07	1.96	.41	1.93	.20	−.01	.35	−.03	.21	.05	.16	.26	28.8	22.0	18.9
Iterative	.67	.12	.69	.07	1.97	.44	1.93	.21	0	.37	−.03	.22	.03	.10	.17	24.4	19.6	16.9
Continuous-updating (.9%)	.64	.15	.67	.08	1.62	10.66	1.94	.24	.21	14.49	−.10	.23	.01	.07	.15	22.0	18.2	16.0
Truncated CU (1.6%)	.64	.14	.67	.08	1.99	.65	1.94	.23	−.08	.63	−.10	.22	.01	.07	.15	21.9	18.0	16.0
ML	.65	.08	.66	.05	1.87	.22	1.86	.13	−.02	.27	−.04	.16						

NOTE: For the continuous-updating GMM approach, figures in parentheses indicate the frequency of samples that were discarded before summary statistics were computed.

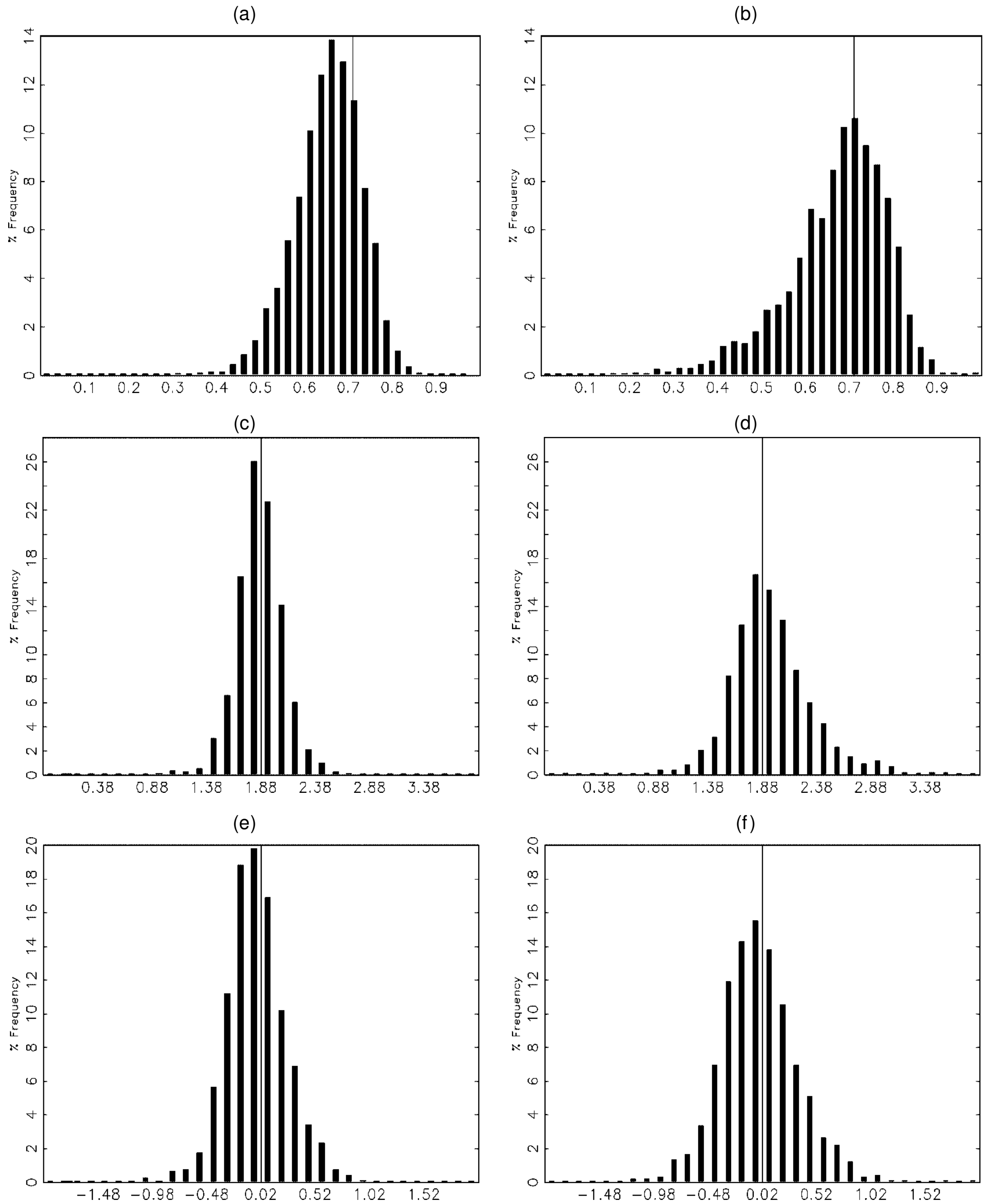

Figure 1. Distribution of Parameter Estimators. (a) ML estimator of ρ; (b) two-step GMM estimator of ρ; (c) ML estimator of β; (d) two-step GMM estimator of β; (e) ML estimator of γ; (f) two-step GMM estimator of γ. Two-step GMM estimators were computed using $\mathbf{S}_{1T}$. The data-generating process is given by (1)–(3), with parameters obtained by ML (Table 2). Sample size is $T = 85$.

which the J statistic exceeds the relevant critical value of the chi-squared distribution, whereas the size-adjusted critical values are defined as the values of the J statistic that are exceeded by the given fraction of the sample J statistic. Four results are worth noting. First, in finite samples, the two-step and iterative GMM tend to reject the overidentifying restrictions too often, whereas the continuous-updating GMM approach has an empirical size closer to the nominal size. This result is consistent with the recommendation of Hansen et al. (1996) and Stock and Wright (2000) to base inference on the continuous-updating criterion. Second, selecting covariance matrices with data-dependent bandwidths ($\mathbf{S}_{2T}$ and $\mathbf{S}_{3T}$) does not improve the finite-sample performances of the J test as compared with the benchmark $\mathbf{S}_{1T}$. In contrast, the estimator $\mathbf{S}_{4T}$ provides correctly sized J statistics, at least for the iterative and continuous-updating GMM. Third, the size-adjusted critical values are substantially larger than the asymptotic critical values taken from the chi-squared distribution. This in particular leads to nonrejection of the overidentifying restrictions, as reported in Table 1. Finaly, unreported investigation reveals that, using the size-adjusted critical values, the J tests have very low power against parameter instability (of the form described in Sec. 5).

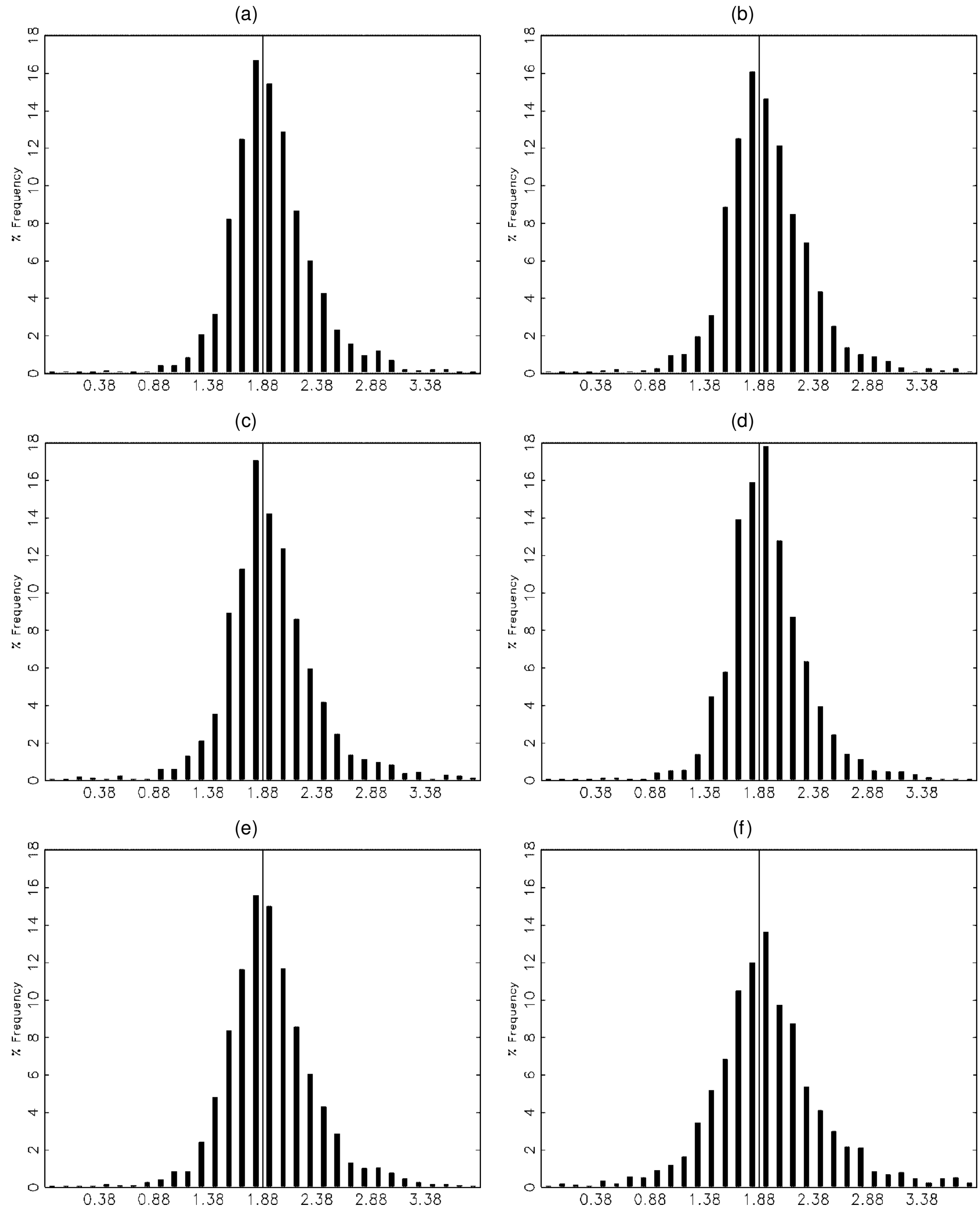

Figure 2. Distribution of Estimators of β. (a) Two-step GMM estimator with $\mathbf{S}_{1T}$; (b) two-step GMM estimator with $\mathbf{S}_{2T}$; (c) two-step GMM estimator with $\mathbf{S}_{3T}$; (d) two-step GMM estimator with $\mathbf{S}_{4T}$; (e) iterative estimator with $\mathbf{S}_{1T}$; (f) continous-updating GMM estimator with $\mathbf{S}_{1T}$.

This can be related to theoretical results pointing to the lack of power of the J test against structral instability (Ghysels and Hall 1990) and explains why misspecification may not be detected in our context.

Overall, our Monte Carlo experiments suggest that (1) the reported finite-sample bias of GMM estimators cannot explain the large discrepancy between estimators; (2) in a small sample like ours, the continuous-updating GMM is very likely to yield unreliable estimators; (3) in most cases, the size-adjusted critical values of the J statistics are so large that the null hypothesis of a correct specification is never rejected; and (4) the efficient covariance matrix $\mathbf{S}_{4T}$ provides correctly sized J statistics.

5. INVESTIGATING PARAMETER STABILITY

Another route for explaining the discrepancy between estimators is misspecification of the model or of the moment conditions. Because the J statistics (reported in Table 1) or the standard residual check (as in Table 2) did not point to any misspecification problem, we now perform an in-depth residuals analysis to identify the source of discrepancy between parameter estimates.

5.1 Subsample Residual Analysis

Figure 3 displays the residuals of the reaction function. The figure suggests two key features in the dynamics of residuals:

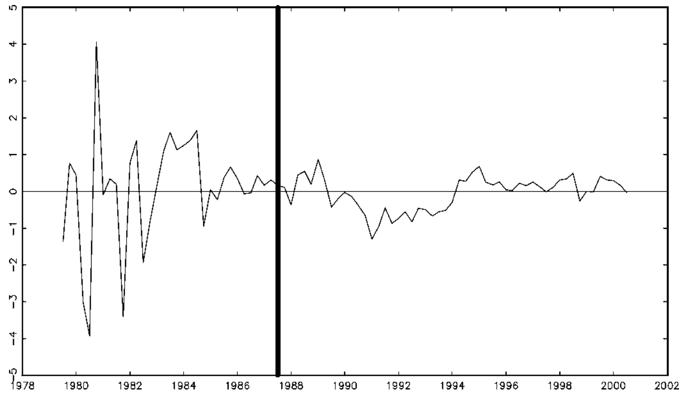

Figure 3. The Reaction Function Residuals Computed From the ML Joint Estimation of (1)–(3).

First, residuals are very volatile over the first part of the sample, and second they are strongly correlated over the second subperiod. To provide further insight into this issue, we focus on the residual properties over the two subsamples. The breaking date is chosen to be 1987:Q3, which corresponds to the change in Federal Reserve chairman.

Summary statistics, reported in Table 4, show that subsample residual properties differ markedly from those obtained over the whole sample. On one hand, whereas serial correlation is barely significant over the whole sample, it is very strongly significant over the 1987–2000 subperiod. It decreases only slowly with the horizon (.74 at horizon 1, and still .38 at horizon 4), a pattern very suggestive of an AR(1) process. On the other hand, whereas heteroscedasticity appears as a key feature of the whole period, it does not turn out to be an issue within each subperiod. A very simple way to describe the variance process would thus be to assume that it is constant with a break at a given date.

5.2 Stability Tests

We now interpret the subsample residual properties as symptomatic of a structural shift in (some of) the reaction function parameters over the sample period. [Note that this does not preclude persistency in the monetary policy shock found by Rudebusch (2002) over the 1987–2000 period. But, as this author pointed out, there is "econometric near-observational equivalence of the partial-adjustment rule and the non-inertial rule with serially correlated shocks" (p. 1164).] This is likely to prevent consistency of ML as well as GMM estimators, because constraining parameters to be constant over the sample period leads to an omitted-variable bias. In addition, the shift in variance also plays an important role, because it prevents the standard diagnostic check from detecting residual serial correlation.

To investigate this issue, we perform parameter stability tests on both ML and GMM estimates. We assume that the shift, if any, occurs in 1987:Q3. Note that we do not suggest that the suspected shift occurs at this date precisely; anecdotal evidence suggests that it may well occur before this date. The period 1979:Q4–1982:Q4, which experienced a change in the operating procedures of the Federal Reserve, had a coincident sharp increase in interest rate volatility. However, the joint decrease in inflation and interest rate lasted up to 1986, suggesting that there may be quite significant uncertainty on the date of the shift. We found it reasonable to treat the break point in 1987:Q3 as known, given the change in chairman as well as the large number of studies focusing on the post-1987 period. It should also be emphasized that selecting such a late break point avoids contaminating the estimation of the reaction function over the end of the period.

Results of stability tests are presented in Table 5. We adopt the strategy developed by Andrews and Fair (1988) for known break point. The test of stability in the ML framework (panel A)

Table 4. Summary Statistics on Sample and Simulated Reaction Function Residuals

	Panel A: Sample data (1979:Q3–2000:Q3)		*Panel B: Simulated data [Data-generating process: eqs. (4) and (5)]*	
	Statistic	*p value*	*Statistic*	*p value*
Full sample				
Q(4)	5.364	.252	5.902	.207
Q(8)	15.973	.043	13.698	.090
R(4)	36.567	.000	17.961	.001
R(8)	47.681	.000	27.651	.001
J–B	125.284	.000	42.731	.000
see	1.016		.878	
First subsample				
Q(4)	3.050	.550	3.678	.451
Q(8)	8.715	.367	7.479	.486
R(4)	10.722	.030	2.951	.566
R(8)	15.751	.046	6.685	.571
J–B	4.923	.085	1.403	.496
see	1.559		1.452	
Second subsample				
Q(4)	75.995	.000	27.152	.000
Q(8)	92.919	.000	34.677	.000
R(4)	8.026	.091	8.082	.089
R(8)	15.457	.051	10.688	.220
J–B	2.605	.272	6.880	.032
see	.460		.417	

Table 5. Stability Tests of the Reaction Function Parameters

	Estimation results			Power of test: Rejection rates using...	
	Statistic	Asymptotic p value	Size-adjusted p value	Asymptotic 5% CV	Size-adjusted 5% CV
Panel A: ML estimates (LR test)					
ML	134.572	.000	.000	.943	.900
Panel B: GMM estimates (Wald test)					
Covariance-matrix estimator $\mathbf{S}_{1T}$					
Two-step GMM	28.668	.000	.754	.978	.150
Iterative GMM	20.028	.001	.877	.995	.194
Covariance-matrix estimator $\mathbf{S}_{2T}$					
Two-step GMM	162.088	.000	.687	.997	.208
Iterative GMM	163.538	.000	.713	.998	.272
Covariance-matrix estimator $\mathbf{S}_{3T}$					
Two-step GMM	182.342	.000	.771	.996	.187
Iterative GMM	208.054	.000	.750	.999	.216
Covariance-matrix estimator $\mathbf{S}_{4T}$					
Two-step GMM	14.507	.013	.700	.818	.110
Iterative GMM	20.706	.001	.548	.845	.199

NOTE: Results for the continuous-updating estimator are not reported, because the algorithm failed to converge over both subperiods.

is performed using a LR statistic. Under the alternative hypothesis, all parameters (including the shock variance) are allowed to shift in 1987:Q3. Following the evidence reported by Judd and Rudebusch (1998), for the 1987–1997 period a second lag of interest rate is introduced in the reaction function, so that under the null, the statistic is asymptotically distributed as a $\chi^2(5)$. Using the asymptotic p values of the statistic, we strongly reject the stability of the reaction function parameters. Given the small size of the subsamples (32 and 53 observations), we also computed size-adjusted p values based on a normal distribution. Our conclusion on the LR test remains unchanged.

Regarding the GMM estimator (panel B), we report the Wald test statistics for the two-step and the iterative GMM for the four long-run covariance matrices. (We do not report results for the continuous-updating GMM because the algorithm failed to converge in many cases, apparently because of the overly small size of subsamples.) Using asymptotic p values, the null is strongly rejected for each GMM estimator. In contrast, the corresponding size-adjusted p values are very large. This suggests that the substantial sample size distortion prevents the estimated test statistics, though very large, from being significant. For instance, with covariance matrices $\mathbf{S}_{2T}$ and $\mathbf{S}_{3T}$, the test statistics exceed 150, so that the asymptotic p values are essentially 0. Yet once the sample-size distortion is taken into account, the p values are larger than 66%, and the null of stability cannot be rejected. Notice that the failure to reject parameter stability does not preclude instability of moment conditions stemming from other sources. This issue has been investigated by Hall and Sen (1999).

To further investigate the discrepancy between ML-based and GMM-based stability tests, we now evaluate the power of the tests for both estimation procedures. We measure this power by simulating 2,000 samples of the model with a shift in 1987:Q3 (with parameter estimates and covariance matrix corresponding to Table 7; see Sec. 5) and computing the fraction of samples for which the null hypothesis is correctly rejected at the 5% significance level. We consider the asymptotic critical values as well as the size-adjusted critical values obtained from the previous simulation experiment (i.e., a simulation of the model with no shift). Interestingly, the ML-based test has a very large power, whatever critical values are used. In contrast, when the sample-size distortion, the power of the GMM-based test is dramatically low, in no case larger than 30%.

In sum, only the ML estimation procedure demonstrates reasonable power against the alternative hypothesis of a break in 1987:Q3. We thus conclude that the stability of the reaction function parameters is strongly rejected. Interestingly, some previous studies using GMM (e.g., Mehra 1999; CGG) tested such a hypothesis but obtained inconclusive evidence; this may be attributed to the lack of power of the GMM-based test of stability.

5.3 Consequences of Heteroscedasticity

An interesting issue raised by our results is that the rejection of stability of the reaction function appears to be very robust for the ML estimation approach. Yet, a standard residual check has been unable to detect such an instability. In fact, the summary statistics on residuals of the model without a shift detected only a significant heteroscedasticity, and no serial correlation.

To illustrate why standard residual check failed to detect serial correlation, consider the following stylized model, with a shift at date τ in parameters as well as in the error variance:

$$y_t = \rho y_{t-1} + \varepsilon_t, \qquad \varepsilon_t \to iid(0, \sigma_1^2), \qquad t < \tau, \tag{4}$$

$$y_t = \rho_1 y_{t-1} + \rho_2 y_{t-2} + \varepsilon_t, \qquad \varepsilon_t \to iid(0, \sigma_2^2), \qquad t \geq \tau. \tag{5}$$

Assuming erroneously that the dynamics of y_t is given by $y_t = \psi y_{t-1} + u_t$ with no shift would yield as residual series

$$\hat{u}_t = \begin{cases} \varepsilon_t + (\rho - \hat{\psi}) y_{t-1} & \text{for } t < \tau \\ \varepsilon_t + (\rho_1 - \hat{\psi}) y_{t-1} + \rho_2 y_{t-2} & \text{for } t \geq \tau. \end{cases} \tag{6}$$

Therefore, omitting the shift in the autoregressive parameters is likely to induce residual serial correlation. This result appears to be inconsistent with our empirical evidence that residuals are not significantly correlated over the whole sample. We argue, however, that our evidence and this stylized model can be reconciled very easily, in finite samples, by allowing a shift in the variance of the policy shock.

For this purpose, we simulated 2,000 samples using (4)–(5) with the parameter estimates and error variance obtained with our data, except that we imposed $\beta = \gamma = \alpha = 0$. Therefore, the experiment was run with $\rho = .564$, $\rho_1 = 1.293$, $\rho_2 = -.492$, $\sigma_1 = 1.488$, $\sigma_2 = .293$, and $\tau = .3$, as in Table 7 (in Sec. 6.2). Then, the model with no shift and one lag is estimated. Summary statistics of this estimation are reported in Table 4 (panel B), so that it can be compared with the statistics of the sample residuals. The key results can be summarized as follows. First, we do not reject the null of no serial correlation over the first subsample, although it appears clear from (6) that residuals should be correlated over the two subsamples. Thus the large variance over the first subsample precludes estimating a significant serial correlation. Second, we do not reject the null of no serial correlation over the whole sample. Here the explanation is that this is the shift in variance (from a high level to a low level of variance) that precludes detecting serial correlation. In contrast, serial correlation is clearly detected within the second subperiod. This appears to match our empirical findings quite closely.

Another concern is that the presence of a strong shift in variance might lead to a spurious break in parameters. To investigate this issue, we simulated the model with no change in parameters ($\rho = \rho_1 = .564$ and $\rho_2 = 0$), while maintaining the assumption of a shift in variance. We tested for stability using a standard Chow test. Using the asymptotic 5% critical value, we found that actual finite-sample size is 4.3% when the variance is constant throughout the sample and .5% when the variance shifts. This result suggests that heteroscedasticity of the form present in the data is very unlikely to create spurious detection of the break, but rather would obscure evidence for a shift.

6. A 1987–2000 REACTION FUNCTION

We consider now the estimation of the forward-looking reaction function over the Greenspan era (1987:Q3–2000:Q3). For GMM, we reestimate (1), while allowing a second lag in the dynamics of the interest rate. For ML, we maintain the estimation of the auxiliary model over the period 1979–2000, because the Phillips curve and the I–S curve have been found to be stable over the whole sample. But we do allow a shift in the reaction function parameters in 1987:Q3, and also allow the covariance matrix of innovations to shift at this date.

6.1 GMM Estimates

Table 6 reports parameter estimates obtained with the different GMM approaches. Several results merit emphasis. First, the autoregressive component of the reaction function differs markedly from that estimated over the 1979–2000 period. This confirms the evidence of Judd and Rudebusch (1998) and rationalizes the serial-correlation pattern found when the reaction

Table 6. GMM Estimates, 1987:Q3–2000:Q3

	Two-step GMM		*Iterative GMM*		*Continuous-updating GMM*	
	Estimate	*SE*	*Estimate*	*SE*	*Estimate*	*SE*
Covariance-matrix estimator $\mathbf{S}_{1T}$						
ρ_1	1.327	.097	1.321	.093	1.305	.067
ρ_2	−.519	.092	−.512	.091	−.504	.064
β	1.661	.337	1.696	.346	1.755	.247
γ	.647	.159	.669	.156	.684	.136
α	1.807	.883	1.741	.907	1.536	.768
J statistic (statistical/asymptotic p value)	*3.361*	*.910*	2.627	.956	2.440	.965
(size-adjusted p value)		.999		.995		.997
Covariance-matrix estimator $\mathbf{S}_{2T}$						
ρ_1	1.336	.074	1.330	.075	1.321	.059
ρ_2	−.530	.080	−.522	.082	−.520	.064
β	1.648	.295	1.686	.303	1.581	.196
γ	.689	.151	.737	.145	.896	.132
α	1.826	.776	1.765	.788	2.243	.626
J statistic (statistical/asymptotic p value)	*7.014*	*.535*	4.060	.852	3.337	.911
(size-adjusted p value)		.998		.974		.998
Covariance-matrix estimator $\mathbf{S}_{3T}$						
ρ_1	1.333	.056	1.310	.051	1.297	.041
ρ_2	−.527	.062	−.494	.059	−.504	.043
β	1.662	.291	1.717	.308	1.897	.261
γ	.647	.154	.784	.153	.698	.110
α	1.807	.847	1.671	.909	1.097	.846
J statistic (statistical/asymptotic p value)	*7.306*	*.504*	5.779	.672	4.232	.836
(size-adjusted p value)		.998		.903		.998
Covariance-matrix estimator $\mathbf{S}_{4T}$						
ρ_1	1.301	.129	1.303	.132	1.296	.075
ρ_2	−.486	.115	−.489	.117	−.493	.064
β	1.723	.387	1.746	.380	1.767	.198
γ	.633	.173	.638	.171	.663	.107
α	1.631	.923	1.561	.903	1.496	.510
J statistic (statistical/asymptotic p value)	*4.820*	*.850*	4.598	.868	2.959	.937
(size-adjusted p value)		.982		.949		.978

Table 7. ML Estimates, 1987:Q3–2000:Q3

Reaction function (1979:Q3–1987:Q2)			Reaction function (1987:Q3–2000:Q3)			Phillips curve			I–S curve		
Parameter	*Estimate*	*SE*	*Parameter*	*Estimate*	*SE*	*Parameter*	*Estimate*	*SE*	*Parameter*	*Estimate*	*SE*
ρ_1	.564	.135	ρ_1	1.293	.094	$\alpha_{\pi 1}$	.402	.110	β_{y1}	1.191	.115
ρ_2			ρ_2	−.492	.079	$\alpha_{\pi 2}$	.011	.087	β_{y2}	−.243	.126
β	1.493	.196	β	1.523	.240	$\alpha_{\pi 3}$	.335	.103	β_r	−.051	.039
γ			γ	.511	.187	$\alpha_{\pi 4}$	.252		β_0	3.434	1.433
α	3.708	.821	α	2.147	.606	α_y	.153	.038			
	Statistic	*p value*		*Statistic*	*p value*		*Statistic*	*p value*		*Statistic*	*p value*
Q(4)	3.044	.550	Q(4)	3.119	.538	Q(4)	*1.424*	*.840*	Q(4)	*2.981*	*.561*
Q(8)	8.386	.397	Q(8)	*10.146*	*.255*	Q(8)	*6.739*	*.565*	Q(8)	*5.210*	*.735*
R(4)	14.521	.006	R(4)	1.443	.837	R(4)	.744	.946	R(4)	*2.966*	*.564*
R(8)	13.733	.318	R(8)	5.286	.948	R(8)	*6.680*	*.878*	R(8)	*6.743*	*.874*
J–B	3.621	.164	J–B	1.893	.388	J–B	1.448	.485	J–B	2.456	.293
see	1.488		see	.293		see	.810		see	.696	
log-L	−230.571										

function is assumed to be stable over the whole sample. Second, point estimates of the parameters obtained by the various GMM procedures are now very close to one another; for instance, estimates of parameter β range between 1.58 and 1.90. Omitting the continuous-updating estimator, the range is even narrower (between 1.64 and 1.75). We obtain a similar result for the other parameters. Third, using different covariance matrices results only in a change in the standard error of parameters, and not in a change in parameter estimates themselves. In most cases, standard errors decrease from covariance matrices $\mathbf{S}_{1T}$ and $\mathbf{S}_{4T}$ to $\mathbf{S}_{2T}$ and $\mathbf{S}_{3T}$.

Last, in no case does the J statistic reject the over identifying restrictions, even when the (less conservative) asymptotic p values are used. In sum, the key insights obtained for the reaction function over the 1987–2000 period are robust to change in GMM options.

6.2 ML Estimates

ML estimates of the model with a shift in the reaction function parameters are presented in Table 7. Note that parameters of the Phillips curve and the I–S curve are not significantly altered by the shift in the reaction function parameters. As far as the reaction function is concerned, we obtain two models with very typical features. Over the first period, corresponding to Volcker's tenure, the inflation parameter is $\beta = 1.49$, whereas the output gap parameter is very close to 0 and insignificant, so that it has been constrained to 0 in the estimate reported in the table. Only one lag of interest rate is significant. Over the second subperiod corresponding to Greenspan's tenure, the reaction function is essentially a dynamic Taylor rule, with $\beta = 1.52$ and $\gamma = .51$. The partial-adjustment model requires a second lag of interest rate to fit the data. These parameter estimates are very similar to those obtained by Rudebusch (2002) over a close sample period.

Statistical properties of residuals are also reported in Table 7. The main features concerning the reaction function residuals are that standard error is divided by 5 between the first and second subperiods, and that the null hypothesis of no serial correlation is not rejected at any significance level for both subperiods, and the heteroscedasticity vanishes over the two subperiods. These results confirm that the serial correlation and the heteroscedasticity obtained over the 1979–2000 sample were to a great extent attributable to the shift in parameters. Figure 4 displays the reaction function residuals when a shift is introduced in (1). It confirms that the reaction function residuals no longer appear serially correlated.

Overall, the discrepancy between GMM and ML estimates is fairly small over the 1987–2000 period. For instance, the baseline estimates of β and γ are 1.65 and .65 for the two-step GMM, compared with 1.52 and .51 for the ML. These estimates differ substantially from those obtained over the whole period, assuming parameter stability. We conclude that the assumptions of a stable output gap parameter and, more importantly, of a stable autoregressive dynamics are responsible for the very large inflation parameter obtained by GMM over the 1979–2000 period. Interestingly, the induced bias has been found to be more pronounced on GMM estimators than on ML estimators.

7. CONCLUSION

In this article we have reexamined the now-standard dynamic forward-looking Taylor rule specification of the Federal Reserve reaction function, implementing both alternative GMM and ML estimation procedures. We have provided some original empirical results. First, over the 1979–2000 period, the

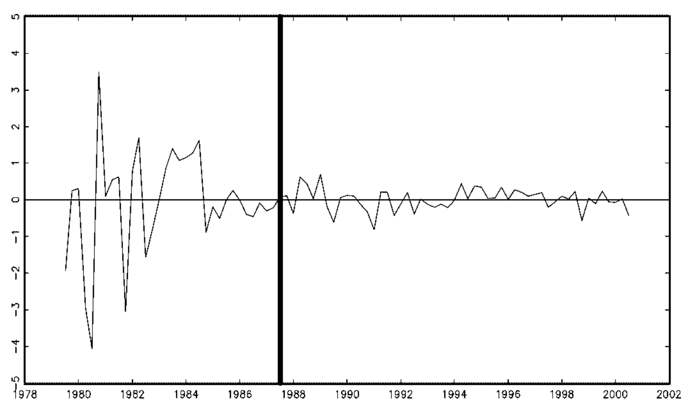

Figure 4. The Reaction Function Residuals Computed From the ML Joint Estimation of (1)–(3), Allowing for a Shift in Reaction Function Parameters in 1987:Q3.

various GMM procedures yield very contrasting estimates. In particular, the continuous-updating GMM, which has not been considered so far in the reaction function literature, produces particularly high and somewhat unrealistic inflation parameters. These results are unlikely to be explained by a finite-sample bias of GMM, which is found to be negligible within our setup.

Second, the ML estimate of the inflation parameter is much lower than GMM estimates, and more in line with the baseline Taylor rule. Analysis of the residuals points only to a strong heteroscedasticity, whereas serial correlation appears to be unimportant. This result seems to be at odds with the evidence of a large discrepancy between parameter estimates. Yet further scrutiny of residuals suggests that misspecification may result from a shift in the reaction function parameters. Indeed, ML-based stability test points to a shift in the reaction function parameters somewhere around 1987:Q3. Surprisingly, however, the shift in the autoregressive parameter does not yield a significant serial correlation of residuals. Our interpretation is that a shift in the policy shock variance obscures evidence of residuals serial correlation over the whole sample.

Third, we showed that over the 1987–2000 period, parameter estimates are very stable across estimation procedures. Moreover, the response to expected inflation is much lower than over the 1979–2000 period, yet consistent with the value suggested initially by Taylor (1993).

In addition to empirical findings, we obtained several results on the properties of estimation procedures for forward-looking monetary policy rules. First, all GMM parameter estimators exhibit large dispersion, although they do not suffer from any economically important finite-sample bias. In contrast, we obtained that the test statistics (J test as well as stability tests) perform very poorly in our context. When the sample size distortion was taken into account, we found that the tests have a very low power. Our overall assessment of GMM in the case of the reaction function is thus less critical than that obtained by Fuhrer et al. (1995) in the case of inventories in terms of parameter bias, but for small samples such as ours, specification tests are very likely to yield unreliable conclusions.

Second, the three GMM estimators considered provide contrasting performances. The two-step and iterative GMM estimators exhibit smaller bias and lower dispersion than the continuous-updating GMM estimator. Moreover, the latter estimator is found to be widely asymmetric and fat-tailed in our setup. This rationalizes the use of a simple approach, as is usually done in empirical studies of the reaction function.

Third, ML is a feasible alternative to GMM for estimating a forward-looking reaction function. A traditional drawback with ML is that it involves estimating a structural model for forcing variables. However, in the present context, a Phillips curve/I–S curve model, such as the Rudebusch–Svensson model, provides a fairly reliable model of the economy. Given the sample sizes typically available for estimating monetary policy rules, ML should be viewed as an attractive alternative to the GMM approach.

ACKNOWLEDGMENTS

The authors thank the editor and three anonymous referees, as well as Patrick Fève, Sylvia Kaufmann, Nour Meddahi, Henri Pagès, and Patrick Sevestre, for their comments. They also thank Simon Gilchrist for providing his AIM algorithm code, and Bob Arnold for providing the CBO data. They also thank seminar participants at the Banque de France, the universities of Evry, Paris 12 Val-de-Marne, Toulouse 1, the T2M conference in Nice, and the ECB workshop on "Monetary Policy Rules" for helpful remarks. The usual disclaimer applies. The Banque de France does not necessarily endorse the views expressed in this article.

[Received July 2001. Revised September 2003.]

REFERENCES

Anderson, G. A., and Moore, G. R. (1985), "A Linear Algebraic Procedure for Solving Perfect Foresight Models," *Economics Letters*, 17, 247–252.

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation-Consistent Matrix Estimation," *Econometrica*, 59, 817–858.

Andrews, D. W. K. (1993), "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Andrews, D. W. K., and Fair, R. C. (1988), "Inference in Nonlinear Econometric Models With Structural Change," *Review of Economic Studies*, 55, 615–640.

Andrews, D. W. K., and Monahan, J. C. (1992), "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 953–966.

Andrews, D. W. K., and Ploberger, W. (1994), "Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative," *Econometrica*, 62, 1383–1414.

Breusch, T., Qian, H., Schmidt, P., and Wyhowki, D. (1999), "Redundancy of Moment Conditions," *Journal of Econometrics*, 91, 89–111.

Burnside, C., and Eichenbaum, M. (1996), "Small-Sample Properties of GMM-Based Wald Tests," *Journal of Business & Economic Statistics*, 14, 294–308.

Clarida, R., Galí, J., and Gertler, M. (1998), "Monetary Policy Rules in Practice: Some International Evidence," *European Economic Review*, 42, 1033–1067.

——— (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics*, 115, 147–180.

Clarida, R., and Gertler, M. (1997), "How the Bundesbank Conducts Monetary Policy," in *Reducing Inflation: Motivation and Strategy* (NBER Studies in Business Cycles, Vol. 30), eds. C. D. Romer and D. H. Romer, Chicago: University of Chicago Press, pp. 363–406.

Cumby, R. E., Huizinga, J., and Obstfeld, M. (1983), "Two-Step Two-Stage Least Squares Estimation in Models With Rational Expectations," *Journal of Econometrics*, 21, 333–355.

Estrella, A., and Fuhrer, J. C. (2003), "Monetary Policy Shifts and the Stability of Monetary Policy Models," *Review of Economics and Statistics*, 85, 94–104.

Ferson, W. E., and Foerster, S. R. (1994), "Finite-Sample Properties of the Generalized Method of Moments in Tests of Conditional Asset Pricing Models," *Journal of Financial Economics*, 36, 29–55.

Fuhrer, J. C., and Moore, G. R. (1995), "Inflation Persistence," *Quarterly Journal of Economics*, 110, 127–160.

Fuhrer, J. C., Moore, G. R., and Schuh, S. D. (1995), "Estimating the Linear Quadratic Inventory Model: Maximum Likelihood Versus Generalized Method of Moments," *Journal of Monetary Economics*, 35, 115–157.

Ghysels, E., and Hall, A. R. (1990), "Are Consumption-Based Intertemporal Capital Asset Pricing Models Structural?" *Journal of Econometrics*, 45, 121–139.

Hall, A. R., and Peixe, F. (2003), "A Consistent Method for the Selection of Relevant Instruments," *Econometrics Review*, 22, 269–287.

Hall, A. R., and Rossana, R. J. (1991), "Estimating the Speed of Adjustment in Partial Adjustment Models," *Journal of Business & Economic Statistics*, 9, 441–453.

Hall, A. R., Rudebusch, G. D., and Wilcox, D. W. (1996), "Judging Instrument Relevance in Instrumental Variables Estimation," *International Economic Review*, 37, 283–298.

Hall, A. R., and Sen, A. (1999), "Structural Stability Testing in Models Estimated by Generalized Method of Moments," *Journal of Business & Economic Statistics*, 17, 335–348.

Hansen, L. P. (1982), "Large-Sample Properties of Generalized Method of Moments Estimator," *Econometrica*, 50, 1029–1054.

Hansen, L. P., Heaton, J., and Yaron, A. (1996), "Finite-Sample Properties of Some Alternative GMM Estimators," *Journal of Business & Economic Statistics*, 14, 262–280.

Hansen, L. P., and Singleton, K. J. (1982), "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models," *Econometrica*, 50, 1269–1286.

Hayashi, F., and Sims, C. A. (1983), "Nearly Efficient Estimation of Time Series Models With Predetermined, but not Exogenous, Instruments," *Econometrica*, 51, 783–798.

Judd, J. P., and Rudebusch, G. D. (1998), "Taylor's Rule and the Fed: 1970–1997," *FRBSF Economic Review*, 3, 3–16.

Ma, A. (2002), "GMM Estimation of the New Phillips Curve," *Economics Letters*, 76, 411–417.

Mehra, Y. P. (1999), "A Forward-Looking Monetary Policy Reaction Function, Federal Reserve Bank of Richmond," *Economic Quarterly*, 85, 33–53.

Nelson, C. R., and Startz, R. (1990), "Some Further Results on the Exact Small-Sample Properties of the Instrumental Variable-Estimator," *Econometrica*, 58, 967–976.

Newey, W. K., and Smith, R. J. (2003), "Higher-Order Properties of GMM and Generalized Empirical Likelihood Estimators," *Econometrica*, forthcoming.

Newey, W. K., and West, K. D. (1987), "A Simple, Positive Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

——— (1994), "Automatic Lag Selection in Covariance Matrix Estimation," *Review of Economic Studies*, 61, 631–653.

Orphanides, A. (2001), "Monetary Policy Rules Based on Real Time Data," *American Economic Review*, 91, 964–985.

Rudebusch, G. D. (2002), "Term Structure Evidence on Interest Rate Smoothing and Monetary Policy Inertia," *Journal of Monetary Economics*, 49, 1161–1187.

Rudebusch, G. D., and Svensson, L. E. O. (1999), "Policy Rules for Inflation Targeting," in *Monetary Policy Rules* (NBER Conference Report series), ed. J. B. Taylor, Chicago: University of Chicago Press, pp. 203–246.

Sack, B., and Wieland, V. (2000), "Interest-Rate Smoothing and Optimal Monetary Policy: A Review of Recent Empirical Evidence," *Journal of Economics and Business*, 52, 205–228.

Sawa, T. (1978), "The Exact Moment of the Least Squares Estimator for the Autoregressive Model," *Journal of Econometrics*, 8, 159–172.

Sims, C. A. (2001), "Stability and Instability in U.S. Monetary Policy Behavior," working paper, Princeton University.

Smith, D. C. (1999), "Finite-Sample Properties of Tests of the Epstein–Zin Asset Pricing Model," *Journal of Econometrics*, 93, 113–148.

Staiger, D., and Stock, J. H. (1997), "Instrumental Variables Regression With Weak Instruments," *Econometrica*, 65, 557–586.

Stock, J. H., and Wright, J. (2000), "GMM With Weak Identification," *Econometrica*, 68, 1055–1096.

Stock, J. H., Wright, J., and Yogo, M. (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, 20, 518–529.

Taylor, J. B. (1993), "Discretion Versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, 195–214.

West, K. D. (1997), "Another Heteroskedasticity- and Autocorrelation-Consistent Covariance Matrix Estimator," *Journal of Econometrics*, 76, 171–191.