Volume 4, Number 4 December 2008

IJCB
INTERNATIONAL JOURNAL OF CENTRAL BANKING
Volume 4, Number 4 December 2008

IJCB

INTERNATIONAL JOURNAL OF CENTRAL BANKING

IJCB International Journal of Central Banking

Volume 4, Number 4 December 2008

The contents of this journal, together with additional materials provided by article authors, are available without charge at ***www.ijcb.org****.*

The *International Journal of Central Banking* is published quarterly (ISSN: 1815-4654). Online access to the publication is available free of charge at **www.ijcb.org**. Individual print subscriptions are available at an annual rate of $100 (USD).

Print subscription orders may be placed online at www.ijcb.org, by phone (+49 69 1344 7623), via fax (+49 69 1344 8553), or by e-mail (editor@ijcb.org).

Renewals, claims, address changes, and requests for permission to reprint material from this journal should be addressed to:

International Journal of Central Banking
DG Research
European Central Bank
Postfach 16 03 19
D-60066 Frankfurt
Germany

Phone: +49 69 1344 7623
Fax: +49 69 1344 8553
E-mail: editor@ijcb.org

ISSN: 1815-4654

Modeling Short-Term Interest Rate Spreads in the Euro Money Market*

Nuno Cassola[a] and Claudio Morana[b]
[a]European Central Bank
[b]Dipartimento di Scienze Economiche e Metodi Quantitativi
Università del Piemonte Orientale, Novara (Italy) and
International Centre for Economic Research (ICER), Torino

In the framework of a new money-market econometric model, we assess the degree of precision achieved by the European Central Bank (ECB) in meeting its operational target for the short-term interest rate and the impact of the U.S. subprime credit crisis on the euro money market during the second half of 2007. This is done in two steps. Firstly, the long-term behavior of interest rates with one-week maturity is investigated by testing for cobreaking and for homogeneity of spreads against the minimum bid rate (MBR, the key policy rate). These tests capture the idea that successful steering of very short-term interest rates is inconsistent with the existence of more than one common trend driving the one-week interest rates and/or with nonstationarity of the spreads among interest rates of the same maturity (or measured against the MBR). Secondly, the impact of several shocks to the spreads (e.g., interest rate expectations, volumes of open-market operations, interest rate volatility, policy interventions, and credit risk) is assessed by jointly modeling their behavior. We show that after August 2007, euro-area commercial banks started paying a premium to participate in the ECB liquidity auctions. This puzzling

*The views expressed in this paper are of the exclusive responsibility of the authors and do not necessarily coincide with those of the ECB. A preliminary version of this paper was presented to the ECB workshop “The Analysis of the Money Market” on November 15, 2007. The authors gratefully acknowledge the comments made by Gabriel Pérez-Quirós and by two anonymous referees. The usual disclaimer applies.

phenomenon can be understood by the interplay between, on the one hand, adverse selection in the interbank market and, on the other hand, the broad range of collateral accepted by the ECB. We also show that after August 2007, the ECB steered the "risk-free" rate close to the policy rate, but has not fully offset the impact of the credit events on other money-market rates.

JEL Codes: C32, E43, E50, E58, G15.

1. Introduction

Like other central banks of major currency areas, the European Central Bank (ECB) implements its monetary policy stance by steering very short-term interest rates. Differently from the U.S. Federal Reserve, the ECB does not announce an explicit target for its operational implementation of the monetary policy stance in the euro area.[1] However, the ECB provides refinancing to the banking system every week through its main refinancing operations (MROs), which are executed via variable-rate tender procedures with a preannounced minimum bid rate (MBR). The level of the MBR is decided by the Governing Council of the ECB at its monthly policy meeting and signals the monetary policy stance for the euro area. Thus, the MBR can be seen as an implicit target for the weekly average of the overnight interest rate. Moreover, the MBR has been set, since April 1999, at the midpoint of the interest rate corridor defined by the rates on the two standing facilities offered by the ECB (marginal lending and deposit facilities). Given that the latter rates have an overnight maturity, they set bounds for the overnight interest rate in the euro area (±100 basis points around the MBR, herein referred to as the interest rate corridor). Still, given that the interest rate corridor is relatively wide, it is unclear how much short-term interest rate volatility the ECB is willing to accept.

[1]For details of the operational framework for monetary policy implementation, see European Central Bank (2006b). For further institutional details of the euro money market, see Ewerhart et al. (2007).

The issue of the efficiency and effectiveness of the implicit, rather than explicit, operational targeting of a short-term interest rate is interesting from the perspective of the evaluation of different frameworks for the implementation of monetary policy and communication policy of the central bank. For instance, is the precision in achieving the target related to it being explicitly announced by the central bank? Another issue is that overnight interbank loans are unsecured, whereas lending by the ECB (and also by the U.S. Federal Reserve) is provided against collateral, which may complicate the quantitative assessment of the precision with which the target is met, as credit and funding/liquidity risk may affect market rates in a time-varying manner. In fact, in the euro area, spreads between very short-term money-market rates and the MBR have varied over time, sometimes displaying higher volatility or even a short-term increasing trend, which led the ECB to enact a policy of providing liquidity above the so-called benchmark amount (see European Central Bank 2002, 2006a). Which factors triggered the widening/narrowing of spreads? Was liquidity policy effective in moving the spreads in the desired direction?

In order to answer the first question, we assess whether the ECB successfully met its implicit operational target for the short-term interest rate. The proposed approach relies on a new fractionally integrated factor vector autoregressive (FI-F-VAR) model, allowing for joint modeling of the common long- and medium-term dynamics of interest rates. The analysis is done in two steps. First, we investigate the source of nonstationarity in short-term money-market rates; second, we test for persistency in the deviations of these rates from the MBR. In the proposed approach, the long-term analysis concerns the source of nonstationarity in the series. Under successful monetary policy implementation, the MBR should determine long-term fluctuations in all short-term interest rates, thereby being the common trend driving the market interest rates. The investigation of the long-term multivariate interest rate structure relies on the use of cobreaking techniques considering the MBR as a deterministic step function. In fact, the MBR is a step function with jumps. From a medium- to longer-term perspective, these jumps are stochastic, as their size and timing are not known in advance: they depend on the business cycle and on the monetary policy "rule" followed by the ECB. However, from a very short-term perspective (e.g., Tuesday on Tuesday) the size and the timing of the jump in the MBR are known in advance, given that the policy announcement is made on a Thursday and the effective implementation is on the following Wednesday.[2] If the common-break process can be identified with the MBR, then, under homogeneity in the cobreaking relationships, the spreads of the various interest rates against the MBR should not be affected by any other source of deterministic nonstationarity. In the second step, the underlying hypothesis is that successful monetary policy implementation should generate stationary and mean-reverting spreads with the tightness of the control inversely related to the degree of persistency in the spreads. Moreover, spreads should be either $I(0)$ or fractionally integrated ($I(d)$, with $0 < d < 0.5$). Hence, the investigation of the medium-term multivariate structure of interest rate spreads relies on the use of (fractional) cointegration techniques. The identified cointegrating vectors have an economic interpretation, which is useful for interpreting the sources of disequilibrium in the money market. The second set of questions is answered by means of forecast-error-variance decomposition and impulse-response analysis where the impact of several shocks (e.g., market expectations, size of operations, overnight rate uncertainty, policy interventions, and credit risk) are tested, also shedding some light on the short- to medium-term determinants of the dynamics of the spreads.

Our main findings are twofold: First, we cannot reject that the MBR is the common trend driving short-term money-market rates, which supports the idea that implicit targeting is an effective communication tool for steering interest rates. Second, we find that spreads against the MBR are highly persistent, which shows that the steering of short-term rates by the ECB is not tight. This may be due (at least partially) to the inefficiency of the implicit targeting, which has to be traded off with other characteristics of the

[2] This is strictly true only after the March 2004 reform of the operational framework of the ECB. In fact, before March 2004 the MBR and the rates of the standing facilities were announced on Thursday and became effective the next day (Friday). This difference should not affect our main results. See European Central Bank (2003, 2005) for details and assessment of the reform.

operational framework.[3] Third, the increasing liquidity deficit in the euro area, expectations of increases in policy rates (or the slope of the yield curve), and volatility of the overnight interest rate put an upward pressure on the spreads; these are effectively counteracted by liquidity policy. Fourth, the long- and medium-term structures of the euro money market do not seem to have been affected by the U.S. subprime credit crisis. However, its short-term dynamics has been significantly disturbed: within the FI-F-VAR model, this can be interpreted as deviations from equilibrium resulting from credit-risk shocks. In this context, we show that after August 2007, the ECB successfully steered the "risk-free" rate close to the MBR, but has not fully offset the impact of the credit events on the other money-market rates.

The remainder of the paper is structured as follows. Section 2 presents the economics of money-market spreads in the euro area. Section 3 presents the econometric methodology. Section 4 reports the main empirical results on money-market spreads. The impact of shocks is presented in section 5. Section 6 concludes.

2. The Economics of Short-Term Interest Rate Spreads in the Euro Area

The *martingale hypothesis* has been the baseline reference for modeling the overnight interest rate in the euro area. It says that when monetary policy is implemented within an interest rate corridor with reserve requirements and averaging, the overnight interest rate on any day within the reserve maintenance period, on_t, should be equal to the rate expected to prevail on the last day of the reserve maintenance period (day T), given the current available information, $E_t(on_T)$. This result can be derived from the seminal work of Poole

[3]Other relevant characteristics are as follows. First, there is the low frequency of open-market operations; weekly operations are less costly than daily operations, namely if the objective of the central bank is to provide refinancing to a large number of banks. Second, there is the size of the reserve requirement; when it is large (above optimal working balances) and remunerated, liquidity shocks do not affect overnight rates on a day-to-day basis because the former are absorbed by variations in the current accounts of the commercial banks with the central bank (except on the last day of the reserve maintenance period). Given the above, and the length of the reserve maintenance period (about one month), short-term interest rates in the euro area are sticky.

(1968) and has been applied to the institutional context of the euro area by Pérez-Quirós and Mendizábal (2006) and Välimäki (2003) (see also Whitesell 2006). Moreover, the expected overnight rate for the last day of the reserve maintenance period should be equal to the probability-weighted average of the rates of the two standing facilities. If policy rates are not changed, the spread of the overnight interest rate against the policy rate can be expressed as follows:

$$\begin{aligned} on_t - MBR &= E_t(on_T) - MBR \\ &= [l \cdot P_t(ML_T)] + [d \cdot P_t(DF_T)] - MBR, \end{aligned} \tag{1}$$

where l is the marginal lending rate, d is the deposit facility rate, and $P_t(.)$ denotes the probability of marginal lending (ML) or recourse to the deposit facility (DF) on the last day of the maintenance period (T), conditional on information available at time t. Equation (1) says that if the reserve maintenance period ends with a liquidity shortage, banks must borrow (overnight) from the lending facility of the ECB at a penalty rate (l); $P_T(ML_T)$ is the probability of such an event. If the reserve maintenance period ends with a liquidity surplus, banks transfer the surplus (overnight) to the deposit facility of the ECB at a penalty rate (d); $P_T(DF_T)$ is the probability of such an event. The latter is equivalent to banks having excess reserves remunerated at a penalty rate.

With a symmetric interest rate corridor, $MBR = (l + d)/2$. If the central bank targets zero net recourse to standing facilities and has unbiased forecasts of aggregate liquidity needs, $P_t(ML_T) = P_t(DF_T) = 0.5$. Substituting the two in equation (1) gives

$$on_t - MBR = \frac{l + d}{2} - MBR = 0, \tag{2}$$

leading to the prediction that the central bank should meet its (implicit or explicit) target without any further action or communication device beyond the following: (i) the announcement of the interest rate corridor and of the MBR, and (ii) the liquidity policy. In practice, a positive spread may still exist due to the unsecured nature of the interbank market or market misperceptions about the liquidity policy of the central bank.

Switching from the overnight interest rate to one-week interest rates, the spread between the one-week EONIA swap rate and the policy rate, $w_t^{swap} - MBR$, is of particular interest. Consider the last week of the reserve maintenance period:

$$w_t^{swap} - MBR = \frac{1}{7}\sum_{s=1}^{7} E_t(on_{t+s}) - MBR = \delta^{swap} + \varepsilon_t^{swap}, \qquad (3)$$

where δ^{swap} is a constant and ε_t^{swap} is a stochastic component, with $E(\varepsilon_t^{swap}) = 0$. From equation (1), by the law of iterated expectations, $\sum_{s=1}^{7} E_t(on_{t+s}) = 7 \cdot E_t(on_T)$, where $T = t + 7$ is the last day in the reserve maintenance period. Substituting in equation (3) gives

$$E_t(on_T) - MBR = \delta^{swap} + \varepsilon_t^{swap}. \qquad (4)$$

The tightness of the ECB's control over the overnight interest rate can be assessed against the benchmark of perfect control, which implies $\delta^{swap} = 0$, and white-noise deviations from target, ε_t^{swap}, $E(\varepsilon_t^{swap}) = 0$, and $E(\varepsilon_i^{swap}, \varepsilon_j^{swap}) = 0$ $(i \neq j)$. In assessing monetary policy implementation, we focus on the persistence and the underlying factors of the deviations of the short-term money-market rate from the target rate. Volatility is only one of those factors.

For other one-week interest rates, w_t^i, we define

$$w_t^i - MBR = \frac{1}{7}\sum_{s=1}^{7} E_t(on_{t+s}) - MBR = \delta^i + \varepsilon_t^i, \qquad (5)$$

where i = *mar*, *war*, *depo*, and *repo* (respectively, marginal and weighted average MRO rates, deposit rate, and repo rate); δ^i is a (small) term premium or discount reflecting the (slightly) longer maturity of the one-week interest rate or differences in credit risk and liquidity risk; and ε_t^i is a stochastic component, with $E(\varepsilon^i) = 0$. Weekly rates are set by market participants in a forward-looking manner. The horizon for expectations considered is the same for all market rates corresponding to the maturity and the frequency of the regular open-market operations of the ECB (MRO). The spreads, $w_t^i - MBR(i \neq swap)$, do not have a clear interpretation; however, the implied spreads among money-market rates have a meaningful economic content. The focus of our analysis is on three spreads that measure credit risk, bid shading, and the (relative) cost of (liquid) collateral.

Credit Risk. Credit risk is captured by the spread between the one-week interbank deposit rate (w_t^{depo}, unsecured lending) and the one-week interbank repo rate (w_t^{repo}, secured or collateralized lending):

$$\begin{aligned} w_t^{depo} - w_t^{repo} &= \left(w_t^{depo} - MBR\right) - \left(w_t^{repo} - MBR\right) \\ &= \delta^{cr} + \varepsilon_t^{cr}, \end{aligned} \tag{6}$$

where $\delta^{cr} = \delta^{depo} - \delta^{repo}$, the difference of the two spreads against the MBR, measures *credit risk*; $\varepsilon_t^{cr} = \varepsilon_t^{depo} - \varepsilon_t^{repo}$, $E(\varepsilon_t^{cr}) = 0$. When the assets of the banks are observed with certainty, δ^{cr} should be close to zero in line with both structural and intensity models of credit-risk pricing (Duffie and Singleton 2003, ch. 3). However, when the assets of the banks are imperfectly observed, namely due to accounting uncertainty, δ^{cr} can be positive even for short maturities (Duffie and Lando 2001; Duffie and Singleton 2003, ch. 5). Moreover, credit risk may vary over the (business) interest rate cycle, increasing when interest rates are increasing and, conversely, decreasing when interest rates are decreasing; these variations should be captured by the residual component, ε_t^{cr}.

Bid Shading. The spread between the one-week EONIA swap rate (w_t^{swap}) and the marginal MRO rate (w_t^{mar}) captures the discount of (marginal) bid rates relative to the unobserved true marginal valuations, the latter proxied by the one-week EONIA swap rate:

$$\begin{aligned} w_t^{swap} - w_t^{mar} &= \left(w_t^{swap} - MBR\right) - \left(w_t^{mar} - MBR\right) \\ &= \delta^{bs} + \varepsilon_t^{bs}, \end{aligned} \tag{7}$$

where $\delta^{bs} = \delta^{swap} - \delta^{mar}$, the difference of the two spreads against the MBR, measures *bid shading*; $\varepsilon_t^{bs} = \varepsilon_t^{swap} - \varepsilon_t^{mar}$, $E(\varepsilon_t^{bs}) = 0$. Bid shading may vary with uncertainty about tender outcomes (i.e., allotment share and tender rates). Unfortunately, no general predictions on the impact of rate uncertainty can be made based on existing multiunit, discriminatory pricing, auction theory. Nevertheless, Ewerhart, Cassola, and Valla (2006) show that for the special case of declining linear marginal valuations and symmetric bidders, in the discriminatory auction the marginal bid rate, the average bid rate, marginal valuations, and bid shading are all expected to increase with a simultaneous and proportional expansion in the liquidity needs of banks and the allotment volume. Hortaçsu (2002), also for some special cases, derives results suggesting that bid shading should increase with rate and liquidity uncertainty and decrease with the number of participants in the auction.

Relative Cost of Liquid Collateral. Consider the spread between the marginal MRO rate (w_t^{mar}) and the one-week repo rate (w_t^{repo}):

$$\begin{aligned} w_t^{mar} - w_t^{repo} &= \big(w_t^{mar} - MBR\big) - \big(w_t^{repo} - MBR\big) \\ &= \delta^{cc} + \varepsilon_t^{cc}, \end{aligned} \tag{8}$$

where $\delta^{cc} = \delta^{mar} - \delta^{repo}$, the difference of the two spreads against the MBR, measures the *relative cost of liquid collateral*; $\varepsilon_t^{cc} = \varepsilon_t^{mar} - \varepsilon_t^{repo}$, $E(\varepsilon_t^{cc}) = 0$. This spread should be positive, $\delta^{cc} > 0$, insofar as commercial banks pledge for central bank operations securities that are less liquid than those accepted for private repo transactions (see Ewerhart, Cassola, and Valla 2006).

3. Econometric Methodology

In our empirical analysis, we jointly model the dynamics of short-term interest rates in levels and test for a single common-break process, for stationarity of spreads against the policy rate (MBR), and for homogeneity of cointegrating relationships among money-market rates. We consider the following fractionally integrated factor vector autoregressive (FI-F-VAR) model (Morana 2007b):

$$x_t = \Lambda_\mu \mu_t + \Xi z_t + \Lambda_f f_t + C(L)(x_{t-1} - \Lambda_\mu \mu_{t-1}) + v_t \tag{9}$$

$$D(L) f_t = \eta_t, \tag{10}$$

where x_t is an n-variate vector of long-memory processes;[4] f_t is an r-variate vector of stationary long-memory factors ($I(d)$, $0 < d_i <$

[4]See Baillie (1996) for an introduction to long-memory processes.

0.5, $i = 1, \ldots, r$); μ_t is an m-variate vector of common deterministic break processes; v_t is an n-variate vector of zero-mean idiosyncratic i.i.d. shocks; η_t is an r-variate vector of common zero-mean i.i.d. shocks; $E[\eta_t v_{is}] = 0$ all i, t, s; Λ_f and Λ_μ are, respectively, $n \times r$ and $n \times m$ matrices of loadings; $C(L) = C_1 L + C_2 L^2 + \ldots$, is a finite-order matrix of polynomials in the lag operator with all the roots outside the unit circle; C_i $i = 1, \ldots$, is a square matrix of coefficients of order n; $D(L) = diag\{(1-L)^{d_1}, (1-L)^{d_2}, \ldots, (1-L)^{d_r}\}$ is a diagonal matrix in the polynomial operator of order r; z_t is a q-variate vector of exogenous/shock variables; and Ξ is an $n \times q$ matrix of parameters. The fractional-differencing parameters d_i, as well as the μ_t and f_t factors, are assumed to be known, although they need to be estimated. This is not going to affect the asymptotic properties of the estimator, since consistent estimation techniques are available for all the parameters and unobserved components.

The Fractional VAR Form. By taking into account the binomial expansion[5] in equation (10) and substituting into equation (9), the infinite-order vector autoregressive representation for the factors f_t and the series x_t can be written as

$$\begin{bmatrix} f_t \\ x_t - \Lambda_\mu \mu_t \end{bmatrix} = \begin{bmatrix} 0 \\ \Xi \end{bmatrix} z_t + \begin{bmatrix} \Phi(L) & 0 \\ \Lambda_f \Phi(L) & C(L) \end{bmatrix} \times \begin{bmatrix} f_{t-1} \\ x_{t-1} - \Lambda_\mu \mu_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{f_t} \\ \varepsilon_{x_t} \end{bmatrix}, \qquad (11)$$

where $D(L) = I - \Phi(L)$, and $\Phi(L) = \Phi_0 L^0 + \Phi_1 L^1 + \Phi_2 L^2 + \ldots$, Φ_i, $\forall i$ is a square matrix of coefficients of dimension r,

$$\begin{bmatrix} \varepsilon_{f_t} \\ \varepsilon_{x_t} \end{bmatrix} = \begin{bmatrix} I \\ \Lambda_f \end{bmatrix} \eta_t + \begin{bmatrix} 0 \\ v_t \end{bmatrix},$$

with variance-covariance matrix

$$E\varepsilon_t \varepsilon_t' = \Sigma_\varepsilon = \begin{bmatrix} \Sigma_\eta{}' & \Sigma_\eta{}' \Lambda_f' \\ \Lambda_f \Sigma_\eta{}' & \Lambda_f \Sigma_\eta{}' \Lambda_f' + \Sigma_v \end{bmatrix},$$

where $E\eta_t \eta_t' = \Sigma_\eta$ and $E v_t v_t' = \Sigma_v$.

[5] $(1-L)^d = \sum_{j=1}^{\infty} \rho_j L^j$, $\rho_j = \frac{\sum_{k=0}^{\infty} \Gamma(j-d) L^j}{\Gamma(j+1)\Gamma(-d)}$, where $\Gamma(\cdot)$ is the gamma function.

Estimation. The estimation problem may be written as follows:

$$\min_{\mu_1,\ldots,\mu_m,f_1,\ldots,f_r,\Lambda_f,\Lambda_\mu,C(L),} T^{-1}\sum_{t=1}^{T}\varepsilon'_{x_t}\varepsilon_{x_t},$$

where $\varepsilon_{x_t} = [I - C(L)L][x_t - \Lambda_\mu\mu_t] - [\Lambda_f\Phi(L)L]f_t - \Xi z_t$. Yet, since the infinite-order representation cannot be handled in estimation, a truncation to a suitable large lag for the polynomial matrix $\Phi(L)$ is required. Hence, $\Phi(L) = \sum_{j=0}^{p}\Phi_j L^j$.

The estimation problem can then be solved following an iterative process, described in Morana (2007b).

The above approach can be understood as a generalization of the factor VAR approach proposed by Stock and Watson (2005), allowing for both deterministic and long-memory stochastic factors. Stock and Watson (2005) provide details about the asymptotic properties—i.e., consistency and asymptotic normality—of the estimation procedure for the case of $I(0)$ variables. Although no theoretical results are currently available for long-memory processes, Monte Carlo evidence provided in Morana (2007a) fully supports the use of the principal component analysis (PCA) for long-memory processes.[6] Moreover, since the fractional-differencing parameter can be consistently estimated, the asymptotic properties of the estimation method are not affected by the conditioning to the initial estimate of the persistence parameter. In addition, the two-step iterated procedure is leading to maximum likelihood estimation of the model and therefore to full efficiency. Finally, the above model can also be estimated by relying on the surplus lag approach, requiring neither the computation of the fractional-differencing parameter nor the binomial expansion, as the $C(L)$ and $\Phi(L)$ matrices can be modeled as standard finite-order stationary polynomials in the lag operator, provided that appropriate accounting of the excess lag is carried out.

4. Data and Modeling Issues

The sample covered in the econometric analysis runs from June 27, 2000, until December 11, 2007. The last eighteen weeks in the sample

[6]Theoretical results also validate the use of PCA in the case of both weakly and strongly dependent processes. See, for instance, Bai (2003, 2004) and Bai and Ng (2004).

(5 percent of the sample) cover a particularly volatile period, after August 2007, related to the impact on the euro money market of the U.S. subprime credit crisis, herein referred to as the *turmoil* period.[7] The following one-week maturity interest rates series are included: the marginal MRO rate, w_t^{mar}; the weighted average MRO rate, w_t^{war}; the uncollateralized loan rate, w_t^{depo}; the collateralized loan rate, w_t^{repo}; and the EONIA swap rate, w_t^{swap}. The data is of weekly frequency, collected on the allotment day of the MRO of the ECB (Tuesday). The market rates were collected at 9:30 a.m. from selected brokers by the Front Office Division (ECB).

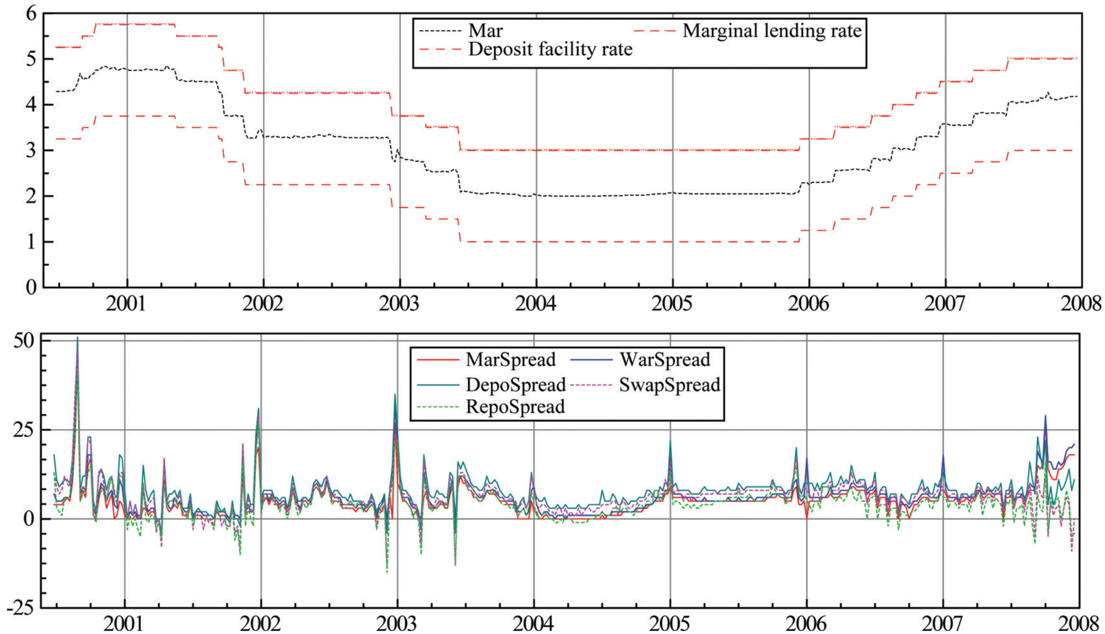
Figure 1 shows, in the upper panel, the time series of the marginal MRO rate and of the rates of the two standing facilities. The former moved smoothly and very close to the midpoint of the interest rate corridor defined by the latter. The lower panel of figure 1 shows the spreads of money-market rates against the MBR. Two facts are noteworthy: first, the relatively smooth behavior of the spreads except for a few spikes mainly around the end of the year; and second, the higher volatility and dispersion after August 2007. The (rounded) sample averages of the spreads, excluding the period between August 2007 and December 2007, are as follows: $w_t^{mar} - MBR = 4$ basis points; $w_t^{war} - MBR = 6$ basis points; $w_t^{depo} - MBR = 8$ basis points; $w_t^{repo} - MBR = 4$ basis points; and $w_t^{swap} - MBR = 6$ basis points.[8] The swap spread is small but different from zero, suggesting less-than-perfect control of the overnight interest rate by the ECB. Still, such a small deviation should not be considered as jeopardizing the monetary policy signaling function of the MBR.

Figure 2 plots the spreads among money-market rates. Credit risk (upper panel) shows some volatility around the end of the year (excluding the turmoil period, the sample average, $w_t^{depo} - w_t^{repo} =$

[7]See European Central Bank (2007) and Ferguson et al. (2007) for an early assessment of the U.S. subprime credit crisis.

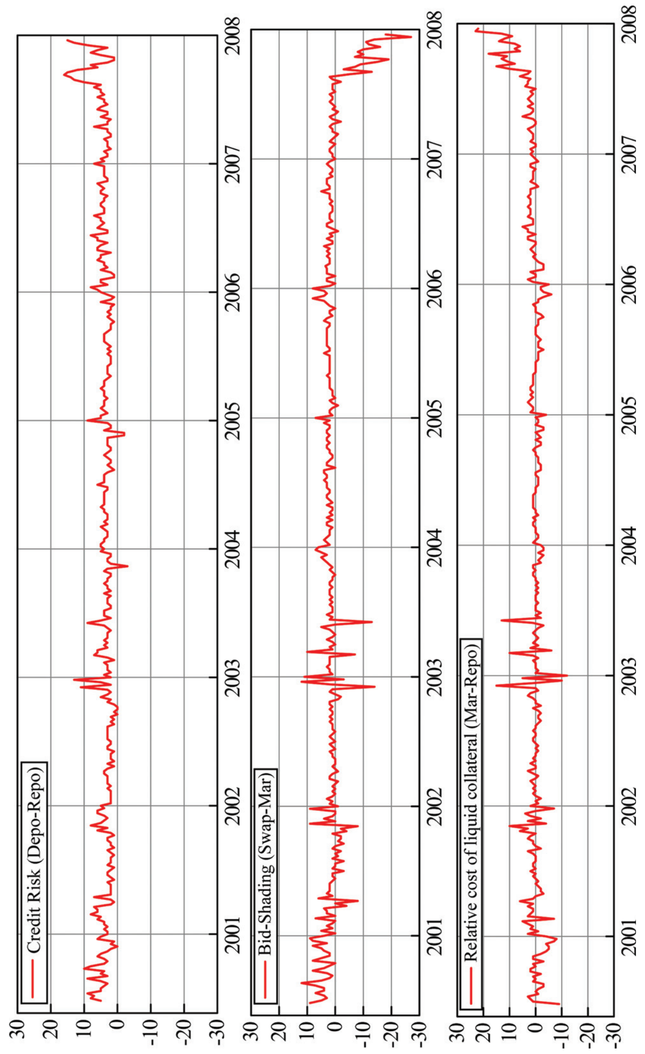
[8]As shown below, the spreads against the MBR follow a stationary long-memory process ($0 < d < 0.5$). Results of Beran (1994, ch. 8) indicate that the sample mean estimator is unbiased and efficient also in the case of long memory. In the Gaussian case the estimator is also the maximum likelihood estimator and therefore it is optimal relative to the class of both linear and nonlinear estimators. The sample mean estimator is also consistent, though the rate of convergence in the stationary long-memory case is slower than for the i.i.d. case: T^{-d} rather than $T^{-0.5}$.

Figure 1. Short-Term Money-Market Rates, Key ECB Policy Rates, and Spreads against the MRO Minimum Bid Rate



Note: Rates are in percent and spreads are in basis points.

Figure 2. Money-Market Rates Spreads (in basis points)



4 basis points). Bid shading (middle panel) is small and shows high volatility around the end of the year (excluding the turmoil period, the sample average, $w_t^{swap} - w_t^{mar} = 2$ basis points). After August 2007, bid shading became negative, which is somewhat puzzling. The relative cost of liquid collateral (lower panel) hovered around zero for most of the sample period, suggesting that scarcity of collateral has not been an issue in the euro area (excluding the turmoil period, the sample average, $w_t^{mar} - w_t^{repo} = 0$). This can be explained by the large pool of collateral accepted by the ECB in its operations.[9] However, during the turmoil period, the relative cost of liquid collateral increased sharply. In fact, the evolution of bid shading and of the costs of liquid collateral, after August 2007, are the two sides of the impact of the turmoil in the euro money market. These developments will be discussed further below.

4.1 Cobreaking Properties and Long-Term Linkages

The multivariate interest rate structure is investigated for cobreaking using the tests of Bierens (2000), with critical values computed by simulation, in order to allow for potential long memory in the break-free series.[10] Two different analyses have been carried out. The first analysis assesses cobreaking between each interest rate series and the minimum bid rate. The second analysis assesses cobreaking among the five money-market interest rate series, jointly considered, with the minimum bid rate. In both cases the cobreaking space has been estimated both unrestrictedly and under the homogeneity restrictions. Moreover, in order to assess the impact of the recent market turmoil, the analysis has been carried out considering two samples. The first sample ends on August 7, 2007, i.e., just before the beginning of the turmoil period (*pre-turmoil* sample, 372 observations); the second sample includes also data since August 14, 2007 (*full* sample, 390 observations). The results of the tests are reported in table 1, panels A and B. As shown in the table, for the pre-turmoil period, both the bivariate and the multivariate analyses point to a single nonlinear deterministic process driving the six

[9] For details on the collateral framework, see European Central Bank (2006b).

[10] See Hendry (1996) for the seminal work on cobreaking and Hendry and Massmann (2007) for a recent survey.

Table 1. Persistence and Copersistence Analysis

A. Bivariate Cobreaking Analysis								
Pre-Turmoil Sample								
#*cbr*	*mar*	*war*	*depo*	*swap*	*repo*	20%	10%	5%
1	0.213	0.179	0.155	0.202	0.087	0.489	0.665	0.803
2	0.951	0.942	0.935	0.935	0.938	0.698	0.876	1.033
1(H)	0.224	0.179	0.155	0.210	0.087	0.489	0.665	0.803
Full Sample								
#*cbr*	*mar*	*war*	*depo*	*swap*	*repo*	20%	10%	5%
1	0.521	0.508	0.241	0.134	0.068	0.489	0.665	0.803
2	0.963	0.924	0.766	0.766	0.774	0.698	0.876	1.033
1(H)	0.660	0.594	0.242	0.134	0.070	0.489	0.665	0.803

B. Joint Cobreaking Analysis					
Pre-Turmoil Sample					
#*cbr*	U	H	20%	10%	5%
1	0.023	0.024	0.489	0.665	0.803
2	0.040	0.043	0.698	0.876	1.033
3	0.055	0.150	0.947	1.075	1.214
4	0.209	0.217	1.084	1.226	1.364
5	0.497	0.937	1.211	1.358	1.466
6	1.401		1.278	1.428	1.515
Full Sample					
#*cbr*	U	H	20%	10%	5%
1	0.038	0.041	0.489	0.665	0.803
2	0.049	0.057	0.698	0.876	1.033
3	0.062	0.133	0.947	1.075	1.214
4	0.224	0.248	1.084	1.226	1.364
5	0.585	1.384	1.211	1.358	1.466
6	1.893		1.278	1.428	1.515

Notes: Panel A reports the value of the Bierens (2000) cobreaking test for each market interest rate (*mar*, *war*, *depo*, *repo*, *swap*) with the minimum bid rate, with (H) and without imposing the homogeneity restriction on the cobreaking space. #*cbr* denotes the number of cobreaking relationships, while 20%, 10%, and 5% are the corresponding critical values, computed by simulation, in order to account for long memory in the candidate break-free series. The null of the test is that the number of cobreaking relationships is less than or equal to n, $n = 1, 2, \ldots, 6$. Similarly, panel B reports the value of the joint Bierens (2000) cobreaking test for all the market interest rates with the minimum bid rate, with (H) and without imposing the homogeneity restriction on the cobreaking space.

(continued)

Table 1. (Continued)

C. Fractional-Differencing Parameter Estimation

Moulines and Soulier (1999) Broadband Log-Periodogram Estimator

Pre-Turmoil Sample

mar	0.311 (0.047)	*war*	0.320 (0.047)	*depo*	0.272 (0.047)	*swap*	0.308 (0.047)	*repo*	0.275 (0.047)

Full Sample

mar	0.295 (0.046)	*war*	0.338 (0.046)	*depo*	0.268 (0.046)	*swap*	0.285 (0.046)	*repo*	0.266 (0.046)

$ARFIMA(4,d,0)$

Pre-Turmoil Sample

mar	0.246 (0.042)	*war*	0.287 (0.044)	*depo*	0.270 (0.044)	*swap*	0.226 (0.043)	*repo*	0.234 (0.043)

Post-Turmoil Sample

mar	0.316 (0.039)	*war*	0.352 (0.040)	*depo*	0.272 (0.042)	*swap*	0.223 (0.042)	*repo*	0.223 (0.043)

D. Fractional Cointegration Analysis (Robinson and Yajima 2002)

Pre-Turmoil Sample

$\#cr$	1	2	3	4	t_1	t_2	t_3	t_4	PV
1%	0.000	0.007	0.026	0.076	0.040	0.080	0.120	0.160	0.952
5%	0.000	0.006	0.022	0.064					
10%	0.000	0.005	0.020	0.058					
20%	0.000	0.004	0.018	0.050					

Full Sample

$\#cr$	1	2	3	4	t_1	t_2	t_3	t_4	PV
1%	0.002	0.008	0.022	0.120	0.040	0.080	0.120	0.160	0.942
5%	0.002	0.007	0.019	0.101					
10%	0.001	0.006	0.020	0.099					
20%	0.001	0.005	0.015	0.080					

Notes: In panel C the estimated fractional-differencing parameters obtained using the Moulines and Soulier (1999) broadband log-periodogram estimator and the ARFIMA(4,d,0) model are reported, with standard errors in parentheses. In panel D the results of the Robinson and Yajima (2002) fractional cointegrating rank test, at the 1%, 5%, 10%, and 20% significance levels, are reported for the selected bandwidth, i.e., four periodogram ordinates. $\#cr$ denotes the number of cointegrating relationships. The null of the test is that the number of cointegrating relationships is less than or equal to n, $n = 1, 2, \ldots, 4$, and t_n denotes the corresponding threshold values. Finally, PV denotes the proportion of total variance explained by the largest eigenvalue. The pre-turmoil period refers to the period June 27, 2000, through August 7, 2007, for a total of 372 observations, while the full sample period refers to the period June 27, 2000, through December 11, 2007, for a total of 390 observations.

series investigated. Concerning the bivariate case, the null of up to a single cobreaking vector (or a single common-break process) is not rejected at the 5 percent significance level, while the null of up to two cobreaking vectors (or structural stability) can be rejected at the 10 percent significance level. Similarly, for the multivariate case, the null of up to five cobreaking vectors (or a single common-break process) driving the six series is not rejected at the 5 percent level, while the null of up to six cobreaking vectors (or structural stability) can be rejected at a significance level slightly higher than 10 percent. In light of the dependence of the money-market interest rates on the MBR, then the single nonlinear deterministic trend detected in the data can be directly associated with the latter series. Since the homogeneity restriction is never rejected at the 5 percent significance level, given the identifying exclusion and normalization restrictions imposed, the cobreaking vectors can be interpreted in terms of spreads against the minimum bid rate in both periods and are irreducible, i.e., of the minimum possible dimension. If the MBR is the only source of structural change in the interest rate series, in light of the above results, the spreads from the minimum bid rate should be purely stochastic. The latter property is actually confirmed by the Dolado, Gonzalo, and Mayoral (2004) structural-break test (DGM test), modified to account for a general and unknown structural-break process (Morana 2007b), pointing to only residual long memory in the break-free series. Similar findings hold for the full sample, suggesting that, conditional on the evidence available so far, the market turmoil appears not to have affected the long-term structure of the euro-area money market.[11]

These results can be used to (partially) answer the first question: the announcement of an MBR is effective for steering short-term money-market interest rates. Moreover, our results suggest that no structural break occurred in the longer-term structure of the money-market spreads since 2000.

[11]For the *pre-turmoil* sample, the p-values of the DGM test, carried out on the spreads against the MBR, are 0.750 for w^{mar}, 0.600 for w^{war}, 0.325 for w^{depo}, 0.530 for w^{swap}, and 0.270 for w^{repo}. On the other hand, figures for the *full* sample are 0.565 for w^{mar}, 0.365 for w^{war}, 0.360 for w^{depo}, 0.495 for w^{swap}, and 0.290 for w^{repo}.

4.2 Persistence Properties of Interest Rate Spreads

The persistence properties of the interest rate spreads are investigated by means of both semiparametric and parametric estimators of the fractional-differencing parameter. Since the data are not affected by observational noise, the broadband log-periodogram estimator of Moulines and Soulier (1999) has been employed. Relative to other semiparametric estimators, it has the advantage of avoiding bandwidth selection problems, being also asymptotically efficient. Moreover, ARFIMA(p,d,q) models also have been fitted to the spreads.[12] The estimates are reported in table 1, panel C, for both the pre-turmoil and full samples. As shown in the table, the degree of long memory is estimated with precision in all cases, pointing to a similar degree of persistence for all spreads. In fact, for the pre-turmoil sample, the estimated fractional-differencing parameter is in the range 0.27 through 0.32 when the broadband log-periodogram estimator is employed and in the range 0.23 through 0.29 when the ARFIMA model is employed. Since in none of the cases the null hypothesis of equality of the fractional-differencing parameter can be rejected at the 5 percent level, the average value of the estimated fractional-differencing parameter has been employed as a common estimate, yielding an overall value of 0.275 (0.045). Moreover, the fractional cointegrating rank test of Robinson and Yajima (2002) points to up to four cointegrating relationships, at the 5 percent significance level, relating the five interest rate spreads to a single common long-memory factor driving the five processes. At the selected bandwidth (four periodogram ordinates), the largest eigenvalue of the spectral matrix at the zero frequency accounts for about 95 percent of total variance, supporting the conclusion drawn in favor of a single common long-memory factor. Similar findings for both the persistence and copersistence analysis hold for the full period, only pointing to a nonsignificant increase in the degree of average persistence, since the estimated fractional-differencing parameter increases to 0.284 (0.044), while the evidence in favor of cointegration is slightly weaker than for the pre-turmoil sample, as the test values are closer to the critical values for the null of cointegration. However, the impact on

[12] An ARFIMA(4,d,0) was selected for all the spreads series on the basis of misspecification criteria. The results are available upon request from the authors.

the proportion of explained total variance is negligible, since for the full sample the largest eigenvalue of the spectral matrix accounts for about 94 percent of total variance.

Concerning the structure of the cointegration space, estimation has been carried out by means of the semiparametric approach of Beltratti and Morana (2006) and Morana (2004a, 2005). The estimated unrestricted vectors are reported in table 2, panel A. Since a single common factor drives the five series, there are four cointegrating vectors, and identification requires a bivariate structure in all cases. As suggested by the theoretical analysis expressed in equation (6) and equation (8), the cointegrating vectors are expressed relative to w_t^{repo}. As the $w_t^{swap} - w_t^{repo}$ spread can be written as the sum of bid shading with the relative cost of collateral, the bid-shading component can be retrieved from the identified cointegrating vectors.

The restricted estimates are reported in table 2, panel A. As shown in the table, in all cases near-homogeneous cointegrating vectors have been found for the pre-turmoil sample. Moreover, taking into account the estimated standard errors, the null of homogeneity cannot be rejected in all cases. The restricted structure is strongly supported by the zero-frequency squared-coherence analysis as well, showing values of the statistic very close to the reference unity value, both in the unrestricted and restricted cases.[13] Finally, support for the proposed identification scheme is also provided by the outcome of the correlation analysis carried out on the restricted and unrestricted factors, pointing to a virtually unitary correlation coefficient between the common factor estimated on the basis of the unrestricted cointegration space and the one obtained on the basis of the restricted cointegration space. The findings are robust to the inclusion of the turmoil data as well, although some differences can be noticed. While the null of homogeneity of the cointegration space cannot be rejected on the basis of the estimated standard errors, a reduction of the efficiency of the estimates, as revealed by the much larger estimated standard errors, can be noticed. Apart from the

[13]The existence of cointegration between $I(d)$ bivariate processes implies that the squared coherence at the zero frequency of the series in differences is equal to 1, while when more than two processes are involved, it is the multiple squared coherence to assume a unitary value. See Morana (2004b).

Table 2. Cointegration Analysis

A. Estimated Cointegrating Vectors									
Pre-Turmoil Sample									
	Unrestricted					**Restricted**			
	CV_1	CV_2	CV_3	CV_4		RCV_1	RCV_2	RCV_3	RCV_4
mar	1.000	0.875	−0.487	0.621	*mar*	1	—	—	—
war	1.034	1.000	0.591	−0.744	*war*	—	1	—	—
depo	−1.153	1.184	1.000	1.314	*depo*	—	—	1	—
swap	0.798	−0.809	0.713	1.000	*swap*	—	—	—	1
repo	0.308	−0.246	0.187	−0.191	*repo*	−0.936	−0.976	−1.081	−1.085
						(0.097)	(0.102)	(0.096)	(0.097)
C^2	0.990	0.993	0.997	0.994	C^2	0.974	0.894	0.985	0.902
Full Sample									
	Unrestricted					**Restricted**			
	CV_1	CV_2	CV_3	CV_4		RCV_1	RCV_2	RCV_3	RCV_4
mar	1.000	0.752	−0.243	0.062	*mar*	1	—	—	—
war	1.031	1.000	0.743	−0.561	*war*	—	1	—	—
depo	−0.267	0.597	1.000	0.999	*depo*	—	—	1	—
swap	0.038	−0.249	0.552	1.000	*swap*	—	—	—	1
repo	0.160	−0.084	−0.019	0.435	*repo*	−1.268	−1.224	−1.212	−0.967
						(0.267)	(0.298)	(0.184)	(0.100)
C^2	0.990	0.994	0.987	0.956	C^2	0.599	0.578	0.755	0.893

(continued)

Table 2. (Continued)

B. Causality Analysis: Bayesian Information Criterion						
Pre-Turmoil Sample						
Predicted Variable	**Predictive Lagged Variables**					
factor	*mar*	*war*	*depo*	*swap*	*repo*	*factor*
BIC	2.538	2.524	2.524	2.535	2.567	2.526
Post-Turmoil Sample						
Predicted Variable	**Predictive Lagged Variables**					
factor	*mar*	*war*	*depo*	*swap*	*repo*	*factor*
BIC	2.365	2.354	2.373	2.431	2.449	2.367

Notes: Panel A reports the estimated unrestricted and restricted cointegrating vectors, with bootstrap standard errors in parentheses. C^2 denotes the squared multiple coherence at the zero frequency. Panel B reports the Bayesian information criterion (BIC) for the predictive equations for the estimated common factor. The pre-turmoil period refers to the period June 27, 2000, through August 7, 2007, for a total of 372 observations, while the full sample period refers to the period June 27, 2000, through December 11, 2007, for a total of 390 observations.

fourth cointegrating vector, in all the other cases there is an increase in the absolute magnitude of the cointegrating parameter. While the increase does not seem to be statistically significant, numerically it is not negligible, i.e., in the range 20–30 percent. The increase in the cointegrating parameter is related to the peculiar behavior of the MRO rates (w_t^{mar} and w_t^{war}), as mean reversion for these rates seems to take longer than for other market rates—a feature that is, however, not inconsistent with the long-memory property detected in the spreads.

An additional difference between the pre-turmoil and full samples can also be noticed in the estimated squared coherence for the restricted case. Yet it would not be correct to argue against fractional cointegration on the basis of the non-negligible reduction in the statistics associated with the exclusion restrictions. For instance, if the bivariate structure was expressed as $w_t^{war} - w_t^{mar}$, $w_t^{depo} - w_t^{war}$, $w_t^{depo} - w_t^{swap}$, $w_t^{swap} - w_t^{repo}$, rather than as spreads against w_t^{repo}, for all rates, figures for the squared coherence would be 0.987, 0.921, 0.794, and 0.893, respectively. Yet a spread structure against w_t^{repo} can be derived from the alternative specification by substitution which is exactly the same as the one estimated directly. Additional supporting evidence can be found by comparing the estimated common long-memory factor in the two cases: the processes are virtually indistinguishable, showing a correlation coefficient close to 0.94, with mean spread equal to -0.03 and 0.08 standard deviation.

Overall, it can then be concluded that, so far, the turmoil has not led to significant changes in the medium-term structure of the euro-area money market. Still, there is some perturbation in the short-term dynamics that will be discussed in the next section. These results can be used to complete the answer to the first question: the persistence found in the spreads suggests less-than-perfect control of short-term interest rates by the ECB (e.g., ε_t^{swap} is not a white-noise process).

4.3 Interpretation of the Factor

Concerning the interpretation of the factor, a Granger causality analysis has been carried out in order to assess whether a stronger predictive power for the common factor could be singled out across the interest rates spreads. In order to control for multicollinearity, the estimated factor has been regressed on its own lags and on the lagged values of each spread, one at a time. As is shown in table 2, panel B, the comparisons of the Bayesian information criteria allows for singling out w_t^{mar} as the only rate whose predictive power is always stronger than of the factor itself. The same result applies to the full sample. These findings suggest that bidding behavior and tender results are the most important determinant of the dynamics of the spreads. These results support the economic interpretation suggested by equations (7) and (8).

Our results highlight the role of the two steps in the implementation of the monetary policy stance by the ECB: the first step sets the level of market rates and consists of the announcement of the MBR (and of the interest rate corridor); the second step consists of steering the spreads against the MBR, by conducting weekly refinancing operations whose results, ultimately, depend on the allotment policy of the ECB.

5. The FI-F-VAR Model

In light of the results of the persistence and copersistence analysis, pointing to a single-break process and a single common long-memory factor, the dimension of the FI-F-VAR model is six equations, corresponding to the five money-market interest rates plus the single common long-memory factor. Given that the common-break process is known—minimum bid rate—the estimation of the model can be performed following a simplified strategy, requiring the iterative procedure only for the joint estimation of the common long-memory factor and the short-term dynamics. To capture the potential economic determinants of credit risk, bidding behavior, and the cost of collateral, five weakly exogenous variables have been included in the specification. The following weakly exogenous variables (z_t) were included:[14]

- F_t is the first difference in the spread between the one-month EONIA swap rate three months forward against the minimum

[14]For all five variables, standard tests do not reject the null of weak exogeneity at the five percent significance level. The p-value of the test ranges between 0.18 and 0.98. Detailed results are available upon request from the authors.

bid rate. This variable is a proxy for changes in interest rate expectations and/or in the slope of the money-market yield curve. Given that it refers to a longer forecast horizon (three months), it should capture the pure effect of expectations beyond the effect of the forward transmission of the very short-term movements in the one-week EONIA swap rate spread. An increase in this variable is expected to put an upward pressure on spreads through the credit-risk component. Before the reform of March 2004, ECB policy rates were implemented within the reserve maintenance period; thus, interest rate expectations also affected the spreads, as they changed the expected levels of the MBR and of the interest rate corridor for the (crucial) last day of the maintenance period, relative to other days. After the March 2004 reform, the impact of interest rate expectations on the spreads should have been much more muted, given that under the new arrangements policy rate changes are implemented only at the start of the reserve maintenance period.[15]

- A_t is the first difference of the residual of an OLS regression of the logarithm of allotment volumes on a constant and a linear trend. Detrending and first differencing is needed because the (log) allotment volume is a nonstationary variable. This variable compares weekly rates of growth in allotment volumes with the trend growth rate. Econometric results are robust to different detrending techniques. This variable captures the effect of shocks to the allotment volumes, and an acceleration of its growth path is expected to put an upward pressure on all spreads against the minimum bid rate (see Ewerhart, Cassola, and Valla 2006, and Välimäki 2006).
- E_t is the squared first difference (Tuesday on Tuesday) in the overnight interest rate (not swap rate). This is a measure of short-term interest rate volatility capturing end-of-maintenance-period conditions and some seasonal factors

[15] In practice, our results should not be affected by these changes because the first half of the sample is dominated by policy rate cuts, and the minimum bid rate put an effective floor on the downward adjustment of tender rates within the maintenance period. This is clearly not the case when policy rates are expected to increase. Nevertheless, testing for a March 2004 structural break in the dynamics of the spreads seems justified.

(end-of-year). It performs better than direct measures of liquidity imbalances as a measure of rate and liquidity uncertainty. An increase in this variable is expected to put an upward pressure on all spreads against the minimum bid rate.

- P_{t-1} is the lagged deviation from benchmark allotment. This variable measures the deliberate, policy-induced deviation from a smooth accumulation of reserve requirements over the reserve maintenance period. An increase in this variable, particularly if persistent and sizable, forces banks to accumulate reserves "too early," thus raising the risk of an early fulfillment of the reserve requirement, which would force banks to park the surpluses at the deposit facility of the ECB, which is costly (100 basis points below the MBR). Thus, an increase in this variable puts downward pressure on all spreads against the MBR.
- CR_t is the iTraxx Financials, which is a credit-risk measure (see Blanco, Brennan, and Marsh 2005); it measures the credit default swap premium on a basket of major European financial firms. It refers to senior debt and has a maturity of five years. Thus, it can be interpreted as the compensation that market participants require in order to bear the default risk of a set of financial firms. Given that it is a market price, it contains the statistical expectation of default risk as well as a risk premium, which is affected by changes in traders' risk appetite. An increase in this variable is expected to put an upward pressure on the spreads that price credit risk.

Coherent with the results of the fractional cointegration analysis, the findings for the pre-turmoil sample point to a single principal component explaining about 94 percent of total variance, as well as between 91 percent and 97 percent of the variance of each of the various interest rate series. Interestingly, the final estimate of the common long-memory factor obtained by means of principal components analysis and the one obtained on the basis of the Gauss/Kasa decomposition, following the approach of Beltratti and Morana (2006) and Morana (2004a, 2005), are strongly correlated (the correlation coefficient is about 0.91) and numerically very close, as the spread between the two factors has zero mean and standard deviation equal to 0.10. A similar finding holds for the full sample, albeit with slightly smaller figures. In fact, the first principal component of the variance-covariance matrix of the spreads explains about 88 percent of total variance, as well as between 84 percent and 96 percent of the variance of each of the various interest rate series. The comparison with the long-memory factor estimated by means of the Gauss/Kasa decomposition also shows a slightly larger deviation, with zero-mean spread, but standard deviation equal to 0.24. Overall, the findings provide further support to the proposed estimation methodology.

Thick estimation (Granger and Jeon 2004) of the FI-F-VAR model has been implemented by allowing up to four lags in the short-memory autoregressive specification ($C(L)$) and twenty-five lags in the long-memory autoregressive specification ($\Phi(L)$),[16] setting Monte Carlo replications to 1,000 for each case, and considering two different orderings of the variables. Hence, the interest rate spread series have been ordered as $w_t^{mar}, w_t^{war}, w_t^{depo}, w_t^{swap}, w_t^{repo}$ in the first case, with the order inverted in the second case. The median estimates have been obtained from cross-sectional distributions counting 8,000 units.[17] While inverting the order of the variables has no consequences for the estimation of the model, it is useful to make the impulse-response analysis and forecast-error-variance decomposition, carried out conditional to a double Choleski identification procedure (see Morana 2007b for details), robust to variable ordering.

5.1 Forecast-Error-Variance Decomposition and Impulse-Response Analysis

As shown in table 3, the results of the forecast-error-variance decomposition are clear-cut. Firstly, the bulk of forecast-error variance for the interest rate spreads is explained by the common shock—interpreted as shocks to MRO outcomes induced by changes in bidding behavior—with the proportion of explained variance increasing

[16] The latter is long enough to describe the long-range dependence in the series. Given the value of the fractional-differencing parameter and the length of the short-term autoregressive specification, the value of the $\Phi(L)$ parameters is negligible beyond the selected truncation order for the binomial expansion.

[17] Detailed results are not reported for reasons of space. A full set of results is available upon request from the authors.

Table 3. Median Forecast-Error-Variance Decomposition

		Pre-Turmoil Sample				**Full Sample**			
		Idiosyncratic			**Common**	**Idiosyncratic**			**Common**
	Horizon (Weeks)	Own	Other	All	—	Own	Other	All	—
mar	1	23.20	0.00	23.20	76.80	28.47	0.00	28.47	71.53
	2	15.71	0.60	16.31	83.69	19.59	0.46	20.05	79.95
	3	15.37	0.72	16.09	83.91	18.90	0.61	19.50	80.50
	4	15.35	0.83	16.18	83.82	18.59	0.66	19.25	80.75
war	1	16.84	1.95	18.79	81.21	22.68	3.87	26.55	73.45
	2	11.35	1.54	12.89	87.11	15.31	3.30	18.761	81.39
	3	10.68	1.67	12.35	87.65	14.62	3.47	18.09	81.91
	4	10.53	1.78	12.31	87.69	14.38	3.60	17.98	82.02
depo	1	11.79	0.31	12.11	87.89	18.54	0.59	18.54	81.46
	2	7.70	0.48	8.18	91.82	12.46	0.62	12.46	87.54
	3	7.25	0.61	7.86	92.14	12.12	0.87	12.12	87.88
	4	7.16	0.66	7.82	92.18	12.01	1.00	12.01	87.99

(*continued*)

Table 3. (Continued)

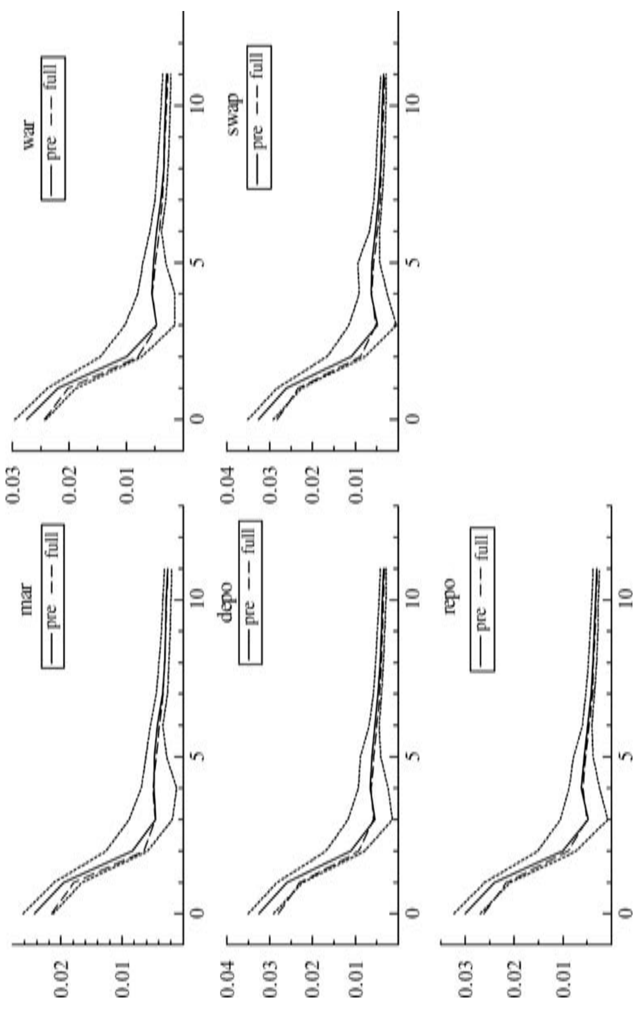
		Pre-Turmoil Sample				Full Sample			
		Idiosyncratic			Common	Idiosyncratic			Common
	Horizon (Weeks)	Own	Other	All	—	Own	Other	All	—
swap	1	13.00	0.33	13.32	86.68	18.60	1.33	19.93	80.07
	2	8.41	0.32	8.72	91.28	12.17	1.26	13.42	86.58
	3	7.90	0.45	8.35	91.65	11.59	1.46	13.06	86.94
	4	7.82	0.54	8.36	91.64	11.42	1.55	12.97	87.03
repo	1	15.84	0.90	16.75	83.25	21.30	3.22	24.52	75.48
	2	10.91	0.91	11.82	88.18	14.62	2.25	16.87	83.13
	3	10.46	0.99	11.44	88.56	13.95	2.25	16.20	83.80
	4	10.46	1.00	11.45	88.55	13.74	2.25	15.89	84.01

Notes: The table reports for each interest rate spread—i.e., the marginal rate (*mar*), the rate (*war*), (*depo*), (*swap*), and (*repo*)—the median forecast-error-variance decomposition, from the one-week to the one-month horizon, obtained from the structural VMA representation of the FI-F-VAR model, following the thick modeling estimation strategy. For each series the table reports the percentage of forecast-error variance attributable to the common and idiosyncratic shocks. For the latter the contribution of the own shock is distinguished from the cumulated effect of the other idiosyncratic shocks. The pre-turmoil period refers to the period June 27, 2000, through August 7, 2007, for a total of 372 observations, while the full sample period refers to the period June 27, 2000, through December 11, 2007, for a total of 390 observations.

with the forecast horizon. In fact, for the pre-turmoil sample, while at the one-week horizon the proportion of variance explained by the common shock is in the range 76 percent to 88 percent, at the one-month horizon the range is 84 percent to 92 percent. Moreover, full stabilization of the effects of the common shock is achieved within one month in all the cases. Secondly, as far as the idiosyncratic shocks are concerned (interpreted as money-market noise), the only non-negligible contribution to the explanation of variance is provided by the own idiosyncratic shock, explaining almost all of the remaining residual variability (in the range 8 percent to 15 percent at the one-month horizon and 12 percent to 23 percent at the one-week horizon), as the other non-own idiosyncratic shocks never explain more than 2 percent jointly. Similar findings can be noted for the full sample, although figures reveal a slightly larger role for the own idiosyncratic shock. In fact, while at the one-week (one-month) horizon the proportion of variance explained by the common shock is in the range 72 percent to 82 percent (81 percent to 88 percent), the proportion of variance explained by the own idiosyncratic shock at the one-week (one-month) horizon is in the range 19 percent to 29 percent (11 percent to 19 percent). The stronger role for the own idiosyncratic shock found for the full sample suggests that the market turmoil has affected the short-term structure of the euro-area money market. This finding is consistent with the cobreaking and cointegration analyses, which on the other hand point to stability in the long- and medium-term structure of the euro-area money market.

Concerning the impulse-response analysis, a similar dynamic reaction to the shocks can be detected for all interest rate series, with only the reaction to the common and the own idiosyncratic shocks being statistically significant. As is shown in figure 3, the dynamic reaction to the common shock points to a hyperbolic decay in all cases, consistent with the long-memory property of the spreads. Yet, the magnitude of the impact of the shock is sizable only within the first two weeks, with a unitary shock leading to a median contemporaneous increase in the spreads close to 3 basis points. Similarly, the decay of the response to the own idiosyncratic shock is also monotonic, though much quicker, and smaller in magnitude, than for the common shock: while the median contemporaneous impact of a unitary own shock is close to 1 basis point, its effects are already halved in magnitude after one week. These findings are consistent

Figure 3. Impulse Responses of Interest Rate Spreads to Common Shock



with the economic interpretation of the common shock (tender outcomes) and idiosyncratic shocks (money-market noise). Findings are very similar, and not statistically different, for both sample periods.[18]

5.2 Impact of Exogenous Variables

Table 4 reports the estimated median contemporaneous impact of the weakly exogenous variables included in the FI-FVAR model, with 95 percent confidence intervals. The upper panel shows the impact before the turmoil. The lower panel shows the impact including the turmoil period. We can use the results to answer the second set of questions asked in the introduction. For the pre-turmoil sample the median contemporaneous impact of the exogenous variables are all statistically significant and have the expected signs. Positive shocks to allotment volumes (A_t), interest rate expectations (F_t), and overnight rate volatility (E_t) put an upward pressure on spreads against the policy rate (MBR). Positive shocks to the policy liquidity variable have a negative impact on the spreads, acting as a counteracting factor (P_{t-1}). As predicted, the increasing liquidity deficit seems to exert a positive impact on bid shading, $w_t^{swap} - w_t^{mar}$, while the effect of volatility on bid shading seems negligible. When the turmoil period is included, a few changes can be noticed, illustrated by the slightly lower estimated impacts for P_{t-1} for all the series, while for A_t and F_t only *mar* and *war* show a smaller impact. Differently, a stronger impact can always be found for E_t. Still, all signs are as predicted. As expected, the impact of credit risk (CR_t) is not statistically significant for w_t^{repo} and w_t^{swap}, because these rates do not price credit risk (see Feldhütter and Lando 2007). The somewhat puzzling finding is that the credit-risk shock had a significant impact on ECB tender rates. In fact, this shock generates two deviations from equilibrium that are two sides of the same coin: an increase in the cost of liquid collateral ($w_t^{mar} - w_t^{repo}$) and a reversion in the sign of bid shading ($w_t^{swap} - w_t^{mar}$). The former is a movement that illustrates the dramatic and sudden disruption in securitization and

[18]Only the impulse-response functions for the common shocks are reported in the plots for reasons of space. A full set of results is available upon request from the authors.

Table 4. Median Impact of Exogenous Variables with 95% Confidence Interval

Pre-Turmoil Sample					
	mar	*war*	*depo*	*swap*	*repo*
P_{t-1}	−0.10	−0.16	−0.24	−0.20	−0.20
	(−0.12 −0.06)	(−0.19 −0.12)	(−0.28 −0.20)	(−0.24 −0.17)	(−0.24 −0.17)
A_t	2.84	3.54	4.85	5.52	4.84
	(2.34 3.29)	(3.04 4.12)	(4.31 5.48)	(4.92 6.44)	(4.35 6.05)
F_t	8.46	9.05	12.70	14.46	11.70
	(6.92 9.57)	(7.51 10.20)	(10.95 14.13)	(12.21 16.71)	(10.17 13.18)
E_t	1.90	3.03	2.35	1.80	1.96
	(0.86 2.80)	(1.58 4.04)	(1.07 3.54)	(0.72 3.17)	(0.79 3.31)
Full Sample					
	mar	*war*	*depo*	*swap*	*repo*
P_{t-1}	−0.04	−0.05	−0.11	−0.11	−0.10
	(−0.08 −0.02)	(−0.12 −0.04)	(−0.24 −0.09)	(−0.21 −0.09)	(−0.19 −0.09)
A_t	2.61	3.38	4.95	5.65	5.02
	(2.00 2.91)	(2.72 3.69)	(4.22 5.29)	(4.90 6.00)	(4.35 5.33)
F_t	7.32	8.24	12.89	14.68	12.08
	(6.30 8.71)	(6.99 9.80)	(11.10 14.62)	(12.65 16.25)	(10.28 13.51)
E_t	2.58	4.16	4.04	2.63	3.10
	(0.15 3.91)	(1.34 5.48)	(0.83 5.48)	(−0.25 4.08)	(0.28 4.41)
CR_t	0.16	0.17	0.13	−0.01	−0.01
	(0.14 0.18)	(0.15 0.19)	(0.12 0.14)	(−0.02 0.01)	(−0.02 0.01)

Notes: The table reports for each interest rate spread—i.e., the marginal rate (*mar*), the rate (*war*), (*depo*), (*swap*), and (*repo*)—the median impact of the exogenous variables (policy (P_t), allotment (A_t), forward (F_t), EONIA volatility (E_t), and credit (CR_t)), with 95% confidence interval. Figures are multiplied by 100. The pre-turmoil period refers to the period June 27, 2000, through August 7, 2007, for a total of 372 observations, while the full sample period refers to the period June 27, 2000, through December 11, 2007, for a total of 390 observations.

the resulting drying up of the asset-backed securities (ABS) market and is interpreted, within our framework, as a short-term disequilibrium movement. The latter illustrates the liquidity premium paid by banks to participate in the ECB tenders, which can be explained by the fact that ABSs are included in the collateral accepted for ECB operations, and have been increasingly used as collateral, while ABSs are not accepted in the private repo market. Moreover, the fact that toward the end of the sample period unsecured money-market rates were below ECB tender rates, $w_t^{depo} - w_t^{war} < 0$, is a clear sign of asymmetric information in the interbank market, pointing to the prevalence of credit rationing during the turmoil.

6. Conclusions

The main conclusions of the paper and the answers to the questions asked in the introduction are as follows. First, one-week interest rates in the euro area are cobreaking and the policy rate is the common-break process; this provides evidence on the effective steering of short-term interest rates by the ECB via the announcement of a minimum bid rate. Second, there is evidence of one common long-memory factor driving interest rate spreads against the policy rate, which is mainly related to shocks to the marginal ECB tender rate spread; this points to bidding behavior and tender outcomes as the driving force behind developments in the money-market spreads against the policy rate. These spreads are conditional on the liquidity policy followed by the ECB over the sample period and, therefore, the mean-reversion properties reveal the preference of the ECB for smoothing short-term interest rate spreads. The detected persistence in the spreads, and their statistical significance, shows that the ECB does not exercise perfect control of short-term money-market rates. This may be related to the low frequency of money-market interventions by the central bank and may characterize operational frameworks like those of the ECB and of the Bank of England which rely on the averaging mechanism of reserve requirements rather than daily open-market operations to stabilize short-term money-market rates (see Nautz and Scheithauer 2008 for a similar conclusion). The results point to another, related source of persistency in the spreads, which is the bidding behavior of counterparties. We found that the conditional means of the spreads are influenced by a number of shocks capturing the impact of changes in interest rate expectations, changes in allotment volumes, and interest rate uncertainty. All these variables put an upward pressure on the spreads. Allotments above benchmark had a downward and stabilizing impact on the spreads, being a counteracting force against the other factors. The market turmoil after August 2007 does not seem to have changed the long- to medium-term structure of the euro money-market spreads. However, credit risk and the associated funding risks have conditioned the short-term dynamics of money-market rates.

References

Bai, J. 2003. "Inferential Theory for Factor Models of Large Dimensions." *Econometrica* 71 (1): 135–71.

———. 2004. "Estimating Cross-Section Common Stochastic Trends in Nonstationary Panel Data." *Journal of Econometrics* 122 (1): 137–83.

Bai, J., and S. Ng. 2004. "A Panic Attack on Unit Roots and Cointegration." *Econometrica* 72 (4): 1127–77.

Baillie, R. T. 1996. "Long Memory Processes and Fractional Integration in Econometrics." *Journal of Econometrics* 73 (1): 5–59.

Beltratti, A., and C. Morana. 2006. "Breaks and Persistency: Macroeconomic Causes of Stock Market Volatility." *Journal of Econometrics* 131 (1–2): 151–77.

Beran, J. 1994. *Statistics for Long-Memory Processes.* New York: Chapman and Hall.

Bierens, H. J. 2000. "Nonparametric Nonlinear Cotrending Analysis, with an Application to Interest and Inflation in the United States." *Journal of Business and Economic Statistics* 18 (3): 323–37.

Blanco, R., S. Brennan, and I. W. Marsh. 2005. "An Empirical Analysis of the Dynamic Relation between Investment-Grade Bonds and Credit Default Swaps." *The Journal of Finance* 60 (5): 2255–81.

Dolado, J. J., J. Gonzalo, and L. Mayoral. 2004. "A Simple Test of Long Memory vs. Structural Breaks in the Time Domain." Mimeo, Universidad Carlos III de Madrid.

Duffie, D., and D. Lando. 2001. "Term Structures of Credit Spreads with Incomplete Accounting Information." *Econometrica* 69 (3): 633–64.

Duffie, D., and K. J. Singleton. 2003. *Credit Risk.* Princeton, NJ: Princeton University Press.

European Central Bank. 2002. "The Liquidity Management of the ECB." *Monthly Bulletin* (May): 41–54.

———. 2003. "Changes to the Eurosystem's Operational Framework for Monetary Policy." *Monthly Bulletin* (August): 41–54.

———. 2005. "Initial Experience with the Changes to the Eurosystem's Operational Framework for Monetary Policy Implementation." *Monthly Bulletin* (February): 65–72.

———. 2006a. "The Eurosystem's Experience with Fine-Tuning Operations at the End of the Reserve Maintenance Period." *Monthly Bulletin* (November): 83–91.

———. 2006b. *The Implementation of Monetary Policy in the Euro Area.* Frankfort: European Central Bank.

———. 2007. *Financial Stability Review.* December.

Ewerhart, C., N. Cassola, S. Ejerskov, and N. Valla. 2007. "Manipulation in Money Markets." *International Journal of Central Banking* 3 (1): 113–48.

Ewerhart, C., N. Cassola, and N. Valla. 2006. "Declining Valuations and Equilibrium Bidding in Central Bank Refinancing Operations." ECB Working Paper No. 668.

Feldhütter, P., and D. Lando. 2007. "Decomposing Swap Spreads." Mimeo, Department of Finance, Copenhagen Business School.

Ferguson, R. W., P. Hartmann, F. Panetta, and R. Portes. 2007. *International Financial Stability.* CEPR and ICMB. Geneva Reports on the World Economy 9.

Granger, C. W. J., and Y. Jeon. 2004. "Thick Modeling." *Economic Modelling* 21 (2): 323–43.

Hendry, D. 1996. "A Theory of Co-Breaking." Mimeo, Oxford University.

Hendry, D. F., and M. Massmann. 2007. "Co-Breaking: Recent Advances and a Synopsis of the Literature." *Journal of Business and Economic Statistics* 25: 33–51.

Hortaçsu, A. 2002. "Bidding Behaviour in Divisible Good Auctions: Theory and Evidence from the Turkish Treasury Auction Market." Mimeo, Stanford University.

Morana, C. 2004a. "Frequency Domain Principal Components Estimation of Fractionally Cointegrated Processes." *Applied Economics Letters* 11 (13): 837–42.

———. 2004b. "Some Frequency Domain Properties of Fractionally Cointegrated Processes." *Applied Economics Letters* 11 (134): 891–94.

———. 2005. "Frequency Domain Principal Components Estimation of Fractionally Cointegrated Processes: Some New Results and an Application to Stock Market Volatility." *Physica A* 335: 165–75.

———. 2007a. "Multivariate Modelling of Long Memory Processes with Common Components." *Computational Statistics and Data Analysis* 52 (2): 919–34.

———. 2007b. "On the Macroeconomic Causes of Exchange Rate Volatility." International Centre for Economic Research, University of Piemonte Orientale.

Moulines, E., and P. Soulier. 1999. "Broadband Log-Periodogram Regression of Time Series with Long Range Dependence." *The Annals of Statistics* 27 (4): 1415–39.

Nautz, D., and J. Scheithauer. 2008. "Monetary Policy Implementation and Overnight Rate Persistence: Some International Evidence." Mimeo, Goethe University Frankfurt.

Pérez-Quirós, G., and H. R. Mendizábal. 2006. "The Daily Market for Funds in Europe: What Has Changed with the EMU?" *Journal of Money, Credit, and Banking* 38 (1): 91–118.

Poole, W. 1968. "Commercial Bank Reserve Management in a Stochastic Model: Implications for Monetary Policy." *Journal of Finance* 23 (5): 769–91.

Robinson, P. M., and Y. Yajima. 2002. "Determination of Cointegrating Rank in Fractional Systems." *Journal of Econometrics* 106 (2): 217–41.

Stock, J. H., and M. W. Watson. 2005. "Implications of Dynamic Factor Models for VAR Analysis." NBER Working Paper No. 11467.

Välimäki, T. 2003. *Central Bank Tenders: Three Essays on Money Market Liquidity Auctions.* Bank of Finland.

———. 2006. "Why the Marginal MRO Rate Exceeds the ECB Policy Rate." Bank of Finland Research Discussion Paper No. 20/2006.

Whitesell, W. 2006. "Interest Rate Corridors and Reserves." *Journal of Monetary Economics* 53 (6): 1177–95.

Monetary Policy under Uncertainty about the Nature of Asset-Price Shocks*

David L. Haugh
Economics Department, OECD
Centre for Applied Macroeconomic Analysis, ANU

The effects of an asset-price movement on inflation and output depend on whether that movement is fundamental or not. However, central banks cannot observe this. This paper examines the issue of how central banks should respond to asset prices given this constraint. Using a modified version of the Gruen, Plumb, and Stone (2005) model, this paper finds it is better to adopt a three-standard-deviation threshold rule for deciding whether to include asset prices in output-gap and inflation forecasts and monetary policy than to ignore asset prices altogether.

JEL Codes: E32, E52, E60.

1. Introduction

There is almost universal agreement in the monetary policy and asset-prices literature that central banks should take account of asset prices (house and equity prices), at least to the extent that they have implications for inflation. A key question that remains, though, is *how* a central bank should incorporate asset-price information in its inflation and output-gap forecasts and monetary policy decisions. One of the judgments central banks face in using asset-price information is deciding whether an asset-price movement is fundamental or not, i.e., justified by a change in the discounted risk-adjusted future stream of returns, or a misalignment or bubble, as the type

*I am particularly grateful to Professor Adrian Pagan, Professor Heather Anderson, and Dr. Mardi Dungey, and also an anonymous referee who provided detailed and useful comments. This paper was mainly written while I was at the Research School of Social Sciences, Australian National University. The views in this paper are those of the author and do not necessarily reflect those of the OECD. Correspondence: 2, rue André-Pascal, 75775 Paris Cedex 16, France; E-mail: David.Haugh@oecd.org.

of movement has different implications for output and inflation and therefore policy.

If there is a fundamental movement, the central bank does not need to adjust policy in response to the asset-price change, because there is no change in the output gap. For example, if asset prices are changing due to a technology shock that causes earnings streams to change, then potential output is changing. In this case, there will be no change in the output gap and no change in inflationary pressure in response to the asset-price change, because any demand increase via channels such as wealth effects or the financial accelerator is matched by an increase in the economy's potential supply.

On the other hand, if the asset-price change is nonfundamental, then potential output will remain unchanged, but demand will shift. This might occur, for example, due to destabilizing speculation and herding in a world of heterogeneous agents (see Delong et al. 1990). This type of movement will result in a change in the output gap and a change in inflationary pressures, which the bank is directly targeting. In this case, it will be necessary for the central bank to adjust policy if it wants to meet its inflation target.

A problem for central banks, though, is that they cannot observe whether an asset-price movement is fundamental or not, and views about the difficulty of judging this directly underpin opinions on the appropriate monetary policy response to asset prices. Bernanke and Gertler (1999) contend it is easier to estimate an output gap than to determine whether an asset-price movement is a bubble. They argue that central banks should ignore stock-price movements that do not appear to be generating inflationary or deflationary pressures. However, Cecchetti, Genberg, and Wadhwani (2002), while acknowledging that it is very difficult to identify small misalignments, argue that central banks can identify very large misalignments—for example, the Japanese stock and property market boom in the late 1980s and the U.S. stock boom in the 1990s—and should take account of these in policy.

One approach to this problem is for the central bank to use a rule to differentiate between asset-price movements and to adjust its forecasts and policy accordingly, rather than treating all asset-price changes in the same manner in policymaking decisions. The central bank never knows with certainty whether an asset-price shift is fundamental, but it does acquire information over time that can assist it in making a decision about whether the change is most likely fundamental or not. One way to use this information is to set a threshold for asset-price changes over which they will be regarded as most likely nonfundamental.

The issues to be examined in this paper include what the appropriate threshold should be and how sensitive this threshold rule is to the assumptions made about the asset-price bubble. To do this, I add a threshold rule to the Gruen, Plumb, and Stone (2005) model and run simulations. The Gruen, Plumb, and Stone model is chosen because it investigates in-depth some of the practical issues in using asset-price information in monetary policymaking. They assume that the central bank can recognize a bubble when it occurs, and they demonstrate that optimal policy depends on what the bank assumes about the nature of the bubble. In this paper I extend their work by assuming that the central bank cannot observe whether a price change is a bubble or not, and I illustrate how the bank's assumption about the nature of the bubble will also affect a threshold rule that the central bank uses to determine what part of an asset-price movement is most likely nonfundamental. The optimal threshold rule is defined as the one that results in the lowest welfare loss in terms of deviations of inflation from the central bank's target and output from potential.

The main findings of the paper include that the optimal threshold is sensitive to the assumptions about the asset-price bubble—in particular, the probability of the bubble bursting and the amount the bubble grows each period. In practice, it is difficult for a central bank to determine these parameter values, but the simulations show that welfare can be improved across almost the entire parameter range by using a threshold of three standard deviations of the asset price rather than by ignoring asset prices. Average welfare across the bubble parameter values can be improved further as the threshold is lowered, but this comes at the expense of a greater range of bubble parameter values over which the chosen threshold rule will lead to a greater loss compared with ignoring asset prices. This suggests that the more risk averse the central bank is, the higher the threshold it should use.

The next section discusses the model used in the simulations and the threshold rule in more detail. Section 3 then discusses the main results and is followed by the conclusion.

2. The Model

The model used in the simulations in this paper is a modified version of the Gruen, Plumb, and Stone (2005) model (GPS model). The GPS model is a two-equation (output gap and inflation), closed-economy model[1] (as per Ball 1999), with the addition of a stochastic asset-price bubble.

2.1 Asset Prices

The market price of an asset, am_t, is assumed to have several components:

$$am_t = af_t + a_t \tag{1}$$

$$= \overline{am_t} + a_t^* + a_t, \tag{2}$$

where af_t is the fundamental asset-price component of the market price. This fundamental component is the sum of $\overline{am_t}$, the long-run average of the market asset price, and a_t^*, the short-run deviation in the fundamental asset price from the long-run average market asset price. The final component of the market asset price is a_t, the deviation of the market asset price from the fundamental asset price, i.e., the bubble or nonfundamental component of the asset price.

The long-run average market asset price is observed by the central bank. The long-run average market price is the anchor for asset prices in the model and is set equal to 0 for convenience. The central bank can also observe the deviation of the market asset price from its long-run average, ap_t, but not its two components:

$$ap_t = am_t - \overline{am_t} \tag{3}$$

$$= a_t^* + a_t, \tag{4}$$

where a_t^* and a_t are the fundamental and nonfundamental parts of the deviation in the market asset price from its long-run average.

[1]The basic conclusions of the simulations will still apply to an open economy. In the open-economy case there would be an extra transmission channel from interest rates to output and inflation via the exchange rate. However, this would not affect the appropriate threshold, which depends on the bubble characteristics rather than the transmission channel for monetary policy.

a_t^* follows an AR(1) process:

$$a_t^* = \tau a_{t-1}^* + \varepsilon_{3t}, \tag{5}$$

where ε_{3t} is a white-noise error term with mean zero and variance σ_3^2. The autoregressive parameter, τ, could be very high so that fundamental deviations from the long-run asset price could be quite persistent.

The bubble, a_t, evolves according to the stochastic process

$$a_t = v_t(a_{t-1} + \gamma), \tag{6}$$

where v_t is a Markov chain and γ is the change in the asset-price level each period. From an asset-price perspective, there are two states of the world: no bubble ($v_t = 0$) and bubble ($v_t = 1$). The bubble can start in any period and can re-form after it has burst. Once the bubble forms, it has a probability p of bursting and probability $1 - p$ of continuing to grow. When $v_t = 0$ the bubble will form with the probability q and not form with probability $1 - q$.

If $v_{t-1} = 1$,

$$\begin{aligned} v_t &= 1 \text{ with probability } 1 - p \\ &= 0 \text{ with probability } p. \end{aligned}$$

Otherwise, if $v_{t-1} = 0$,

$$\begin{aligned} v_t &= 1 \text{ with probability } q \\ &= 0 \text{ with probability } 1 - q. \end{aligned}$$

The unconditional probabilities of each state are given by

$$\begin{aligned} pr(v_t = 1) &= \frac{1 - (1 - q)}{2 - (1 - q) - (1 - p)} \\ &= \frac{q}{q + p} \end{aligned} \tag{7}$$

and

$$pr(v_t = 0) = \frac{p}{p + q}. \tag{8}$$

2.2 Output and Inflation

The output gap is determined by

$$y_t = -\beta r_{t-1} + \lambda y_{t-1} + \iota a_t + \varepsilon_{1t}, \tag{9}$$

where y_t is the output gap, r_t is the deviation of the real interest rate from neutral,[2] ι is the fraction of wealth spent in the current period, a_t measures the size of the asset-price bubble, and ε_{1t} is a white-noise error term.

A key difference between this model and the GPS model is that ιa_t ($\iota < 1$) is added to (9) rather than adding Δa_t to (9). A fraction of the level of a_t rather than Δa_t is added to introduce different properties for the connection between asset prices and the output gap.

The GPS approach of adding Δa_t has a number of implications. First, consumption smoothing occurs entirely via the lag in y_t in (9), as the full amount of the change in the asset price flows through into the output gap. Second, if the bubble were to stop growing, but not collapse, then the output gap would return to zero even though the bubble was still in existence. Third, the economy will operate with a large amount of excess supply when the bubble bursts, as the full value of the bubble will be subtracted from the output gap. Effectively, the economy adjusts its spending in full and immediately to the bursting of the bubble.

By contrast, the model in this paper, which adds a fraction of the level of a_t, introduces an extra layer of consumption smoothing on top of the effect of the lag of y_t in (9), at least as the bubble forms. Adding the level also means the bubble could continue to have an effect on the output gap even if it stopped growing but did not burst. If the bubble bursts in the main simulations, then the positive effect of asset prices on the output gap will cease and the positive output gap will eventually reduce to zero through the term

---

[2]A simplification used here is that the central bank sets the real interest rate. In reality, it sets the nominal interest rate and, unlike the situation in this model, cannot always set the real interest rate as far below neutral as it wants, because it faces a zero lower bound for nominal interest rates. Robinson and Stone (2005) extend the Gruen, Plumb, and Stone (2005) model by incorporating a nominal interest rate and a zero lower bound. They find that this does impose a further constraint on policymaking, but it does not materially alter the basic conclusions of the Gruen, Plumb, and Stone model.

λy_{t-1}. However, it is possible that the bursting of a bubble could have more severe negative consequences for the real economy. In particular, after bursting, the bubble may overshoot its fundamental value, leading to a negative bubble, which pushes output below equilibrium. To explore this possibility and test the robustness of the main findings of the paper to the possibility of negative bubbles, further simulations are conducted in section 3.

Excess demand is expected to affect inflation with a lag. Inflation is determined by

$$\pi_t = \pi_{t-1} + \alpha y_{t-1} + \varepsilon_{2t}, \tag{10}$$

where π_t is the deviation of inflation from target and ε_{2t} is a white-noise error term.

The model is parameterized by Ball (1999) for the United States. The parameter values are $\alpha = 0.4$, $\beta = 1$, and $\lambda = 0.8$. The parameter values imply that each period in the model is a year in length. ι is set equal to 0.03, which represents the consensus view of the marginal propensity to consume out of wealth in the United States (Poterba 2000).

2.3 Monetary Policy

The central bank is assumed to be a flexible inflation targeter; i.e., its objective is to stabilize output around potential and inflation around its target. The central bank sets the interest rate to minimize the following loss function:

$$L_t = \sum_{t=0}^{\infty} \delta^t \left[E_t\left(y_{t+1}^2\right) + \mu E_t\left(\pi_{t+1}^2\right)\right], \tag{11}$$

where the central bank discounts future welfare by δ and μ is the weight on future deviations in inflation from the bank's target rate. The target inflation rate is assumed to be zero for convenience.

To attempt to achieve this objective, the central bank must have information on the output gap and inflation. As it can never observe the actual output gap or the bubble, it estimates the output gap and inflation according to

$$yf_t = -\beta r_{t-1} + \lambda yf_{t-1} + \kappa_t \iota ap_t \tag{12}$$

$$\pi f_t = \pi_{t-1} + \alpha yf_{t-1}, \tag{13}$$

where yf_t and πf_t are the central bank's estimates of the output gap and inflation, respectively. κ_t is an indicator variable:

$$\begin{aligned}\kappa_t &= 1(ap_t > \psi\sigma_{ap_t})\\ &= 0 \text{ otherwise,}\end{aligned}$$

where σ_{ap_t} is the standard deviation of the difference between the long-run and market asset prices and ψ is the threshold number of standard deviations chosen by the central bank. The central bank uses σ_{ap_t} in its threshold rule because, unlike the bubble, it can observe this quantity.

The central bank takes account of asset prices in its output-gap forecasts in a nonlinear way. If the deviation of the asset price from the long-run asset price is greater than ψ standard deviations, κ_t is set to unity, and the central bank judges that the asset-price shift is most likely a bubble and relevant to the output-gap forecast. It then includes a fraction of the deviation from the long-run asset price, ap_t, in its forecast of the output gap.[3]

The optimal interest rate rule for a central bank that can identify bubbles and takes account of them, and where the bubble is independent of monetary policy, is a function of the variables in the system and the expected effect of asset prices on the output gap in the next period, given the state of the bubble in this period.

[3]Rather than using a threshold rule, the central bank could also make inferences about the true output gap and whether there is an asset-price bubble by comparing its forecast of inflation from equation (13) with the inflation outturn. If the bank repeatedly observed that actual inflation was above forecast, this would suggest that the actual output gap was higher than estimated, indicating the presence of a bubble, as a repeated error would be unlikely to arise from shocks to inflation, ε_{2_t}. Simulations were run comparing the threshold rule with an inflation-errors rule where the central bank would consider the asset-price movement a bubble when actual inflation exceeded forecast inflation for five years. The results showed that the three-standard-deviation threshold rule generated lower welfare losses across a much wider parameter range for the bubble than an inflation-errors rule. The parameter values where the welfare loss was lower using the threshold rule were identical to those where the threshold rule had lower welfare losses ignoring asset prices, except where $p = 0.1$, where the inflation rule generated lower welfare losses. This result is perhaps not surprising because the inflation-errors test involves making a more indirect inference about whether there is an asset-price bubble than the inference made from observing total asset-price changes.

A standard result for linear-quadratic[4] dynamic programming problems such as finding the optimal monetary policy in this model is that the optimal interest rate will be a linear function of the variables in the system and any known series that affects the system, in this case the expected effect of a_t on the output gap (see Chow 1973, Bertsekas 2000, and appendix 1). An assumption of this solution method is that asset prices are exogenous in the model (there is no feedback from the economy or interest rates to asset prices) and so they can be treated as predetermined in the dynamic programming problem. Gruen, Plumb, and Stone (2005) also show that the difference between the standard optimal interest rate rule used by a central bank that ignores asset-price movements and the one used by a central bank that takes account of them is the expected effect of the asset-price bubble on output in the next period.

Interest rates affect output with a lag, so to optimally take account of asset prices, the central bank sets interest rates in this period so they will offset the expected effect of asset prices on the output gap in the next period. The expected effect of asset prices on the output gap next period depends on the way in which asset prices are assumed to affect the output gap in (9). If Δa_t is used, as in the GPS model, then the optimal rule is

$$r_t = \phi_1 y_t + \phi_2 \pi_t + \phi_3 (v_t((1-p)\gamma - pa_t) + (1-v_t)(q\gamma)). \tag{14}$$

Because the output gap is a function of Δa_t, the expected effect of asset prices on the output gap next period, given the bubble state, is equal to $E(\Delta a_{t+1}|v_t)$. If $v_t = 1$ (i.e., there is a bubble), then with probability $(1-p)$ the bubble will still be in existence next period and Δa_{t+1} will equal γ; otherwise, with probability p the bubble will burst and Δa_{t+1} will equal $-a_t$. The expected effect of asset prices on the output gap next period, given a bubble this period, $E(\Delta a_{t+1}|v_t = 1)$, is equal to $(1-p)\gamma - pa_t$.

If $v_t = 0$, then with probability q a bubble will form and Δa_{t+1} will equal γ (the bubble will be γ in size in the first period of its existence), and with probability $1-q$ no bubble will form and Δa_{t+1} will equal 0. The expected effect of asset prices on the output gap

[4]The system of equations is linear and the objective function is quadratic in variables that are in the system.

next period, given no bubble this period, $E(\Delta a_{t+1}|v_t = 0)$, is equal to $q\gamma$.

However, if a_t affects the output gap as in the modified model I use in this paper, then the rule is

$$r_t = \phi_1 y_t + \phi_2 \pi_t + \phi_3(v_t \iota(1-p)(a_t+\gamma) + (1-v_t)((\iota q\gamma)). \quad (15)$$

The output gap is a function of a_t, so the expected effect of asset prices on the output gap next period, given the bubble state, is equal to $E(a_{t+1}|v_t)$. If $v_t = 1$, then with probability $(1-p)$ the bubble will still be in existence next period and a_{t+1} will equal $a_t + \gamma$, and with probability p the bubble will burst and a_{t+1} will equal 0. The expected effect of asset prices on the output gap next period, given a bubble this period, $E(a_{t+1}|v_t = 1)$, is equal to $\iota(1-p)((a_t+\gamma))$ (multiply by the fraction ι because only a fraction of wealth is consumed in each period).

If $v_t = 0$, then with probability q a bubble will form and a_{t+1} will equal γ, and with probability $1 - q$ no bubble will form and a_{t+1} will equal 0. Given no bubble this period, the expected effect of asset prices on the output gap next period will be $\iota q\gamma$.

In terms of the monetary policy framework in the paper, ϕ_1, ϕ_2, and ϕ_3 are the weights on the information included in the rule. In the simulation exercise in this paper, these weights are chosen optimally by using the dynamic programming method to solve the linear-quadratic problem for the coefficients in the monetary policy rule (17). Further details are provided in appendix 1.[5]

Which of Δa_t or a_t is added to the output-gap equation will affect the optimal policy as the bubble builds. Adding a_t, as is done in the model in this paper, has the implication that the central bank will increase interest rates above neutral in response to a bubble. In contrast, when adding Δa_t, as is done in the GPS model, there is a point where the expected negative effect of the bubble bursting outweighs its positive effects and the central bank will want to reduce interest rates below neutral. Interest rates will be cut sharply in both models if the bubble bursts.

[5] ϕ_1 and ϕ_2 are obtained from vector G in appendix 1, while ϕ_3 is from vector g in appendix 1. $(v_t((1-p)\gamma - pa_t) + (1-v_t)(q\gamma))$ or $(v_t\iota(1-p)(a_t+\gamma)+(1-v_t)((\iota q\gamma))$ are equivalent to b_t in appendix 1. For the model parameters used here, the optimal weights are $\phi_1 = 1.09$, $\phi_2 = 0.72$, and $\phi_3 = 1$.

Figure 1 shows the results of simulating both the GPS (Δa_t) model and the new model in this paper (a_t) for ten periods, starting in period 1, with bubbles in each model scaled so that the effect of the asset-price bubble shock on the output gap y_t is of a similar magnitude. A bubble commences in period 1 and builds to a peak in period 5 (marked by the vertical dotted lines) and then bursts in period 6. The GPS results are in the left panels. The top left panel shows the deviation of the real interest rate from neutral, r_t, with the black line showing the interest rate set by a central bank that can identify bubbles and which takes account of their expected effect next period according to (14) (the activist). The grey line in the top left panel shows the policy of a central bank that expects no asset-price effect on the output gap next period (the skeptic) and sets their interest rate according to

$$r_t = \phi_1 y_t + \phi_2 \pi_t. \tag{16}$$

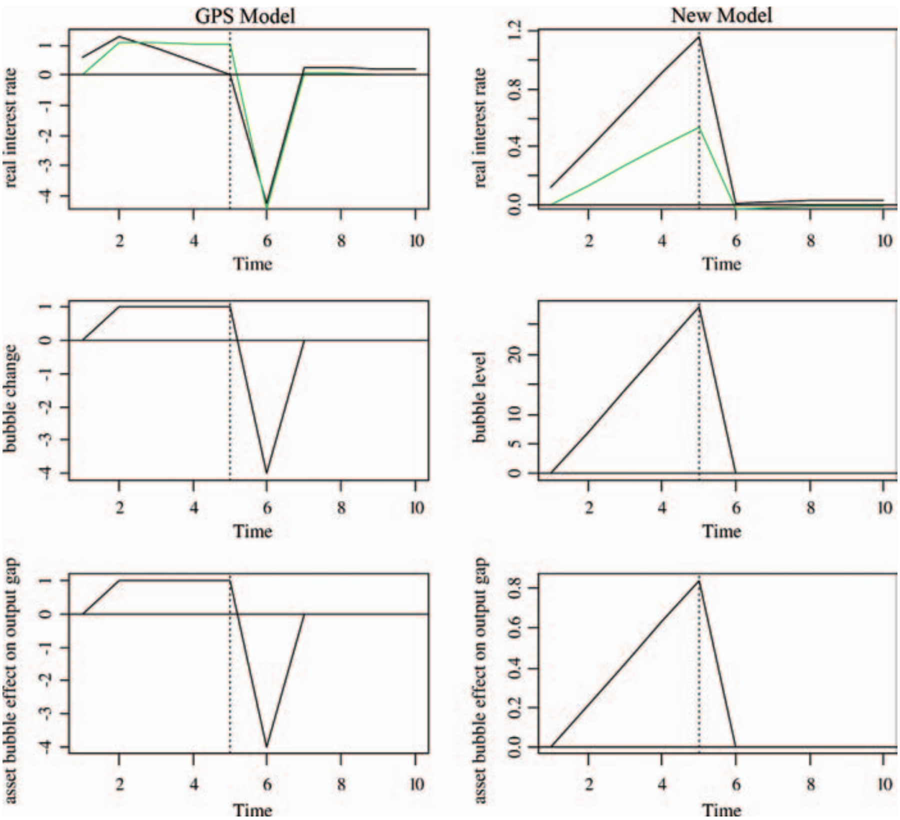
The policy recommendations implemented in the economy are those of the skeptic central bank. The top right panel shows the equivalent results for the new model in this paper, where a_t is added to the output-gap equation, with activist policy being set according to (15). The middle left panel shows Δa_t, which is added to the output-gap equation in the GPS model, while the middle right panel shows a_t. The bottom panels show the effect of the asset-price bubble on the output gap in each model.[6]

In the GPS model, the activist central bank initially raises interest rates above neutral in response to the bubble, but then begins to moderate this tightening from period 2 onward as the expected negative effect of the bubble bursting begins to counter its ongoing positive effects. When the bubble bursts in period 6, interest rates are cut sharply. In the level (a_t) model in the right panel, the activist continues to tighten interest rates as the positive effect on the output gap from the bubble continues to grow. When the bubble bursts, the activist central bank also cuts interest rates sharply.

---

[6]In the GPS model, Δa_t is equivalent to the real shock to the economy from asset prices, while in the new model a_t is a component of asset prices, which has an effect on the real economy via wealth effects and depends on the fraction of wealth spent in each period, ι. To make the two shocks to the real economy similar in each simulation, γ_t is set equal to 1 in the GPS model and 7 in the new model.

Figure 1. Simulating the Effect of a Bubble on Monetary Policy



The optimal rules can be used by the central bank if it can identify bubbles. However, it is assumed in the simulations (except the base one) that this is not the case. Instead, the central bank uses the information that it does know, viz. the entire deviation from the long-run asset price, ap_t, and adopts the rule given by

$$rf_t = \phi_1 yf_t + \phi_2 \pi f_t + \kappa_t \phi_3 \iota (1 - pf_t)(ap_t + \gamma f_t), \qquad (17)$$

where pf_t is the bank's estimate of p and γf_t is the bank's estimate of γ. As described above, if the market asset-price deviation from the long-run asset price is more than a certain threshold number of standard deviations, then the central bank will regard it as most likely a bubble, i.e., $\kappa_t = 1$. In this case, the bank will be concerned about the effects on inflation and will want to take account of its estimate of the expected effect of the asset-price bubble on the output gap next period, $(1 - pf_t)\iota ap_t + \gamma f_t$. The bank estimates the probability of the bubble bursting, pf_t, from the average length of periods it has regarded as bubbles in the last 100 years:[7]

$$pf_t = \frac{1}{\overline{bl}_t}, \tag{18}$$

where $\overline{bl}_t$ is the average bubble length in the last 100 years.

The bank estimates the size of the change in the bubble from the last change in ap_t:

$$\gamma f_t = ap_t - ap_{t-1}. \tag{19}$$

The bank will only use this estimate when it considers a bubble to be in existence and most of the movement in ap_t to be due to the bubble.

2.4 The Threshold Rule

The central bank uses a threshold rule for deciding whether an asset price is most likely a bubble because it cannot observe whether an asset-price movement is a bubble or not. The threshold rule is used on the grounds that if the asset-price shift is sufficiently distant from the long-run asset price, it is unlikely to represent a fundamental shift and therefore is most likely a bubble.

The deviation of the market price from the long-run asset price, ap_t, is the sum of a_t^*, the short-run deviation in the fundamental asset price from the long run, and a_t, the bubble, so the variance of ap_t used in the threshold rule is a function of the variance of both a_t^* and a_t.

[7]Results from simulations where the central bank uses only the last twenty-five years of asset-price data to calculate the probability of the bubble bursting and the variance of ap_t (see further on in the main text) were similar to the full simulation results. In particular, the main result that using the three-standard-deviation rule performs better than ignoring asset prices across most of the parameter range for variables that affect the variance of a_t relative to ap_t remains.

The mean and variance of ap_t, a_t^*, and a_t are given by the following (for details see appendix 2):

$$E(ap_t) = E\left(a_t^*\right) + E(a_t), \tag{20}$$

where

$$E\left(a_t^*\right) = 0 \tag{21}$$

$$E(a_t) = \frac{q\gamma}{(q+p)p} \tag{22}$$

$$var(ap_t) = var\left(a_t^*\right) + var(a_t)\left(a_t^* \text{ is independent of } a_t\right), \tag{23}$$

where

$$var\left(a_t^*\right) = \sigma^2/(1-\tau^2) \tag{24}$$

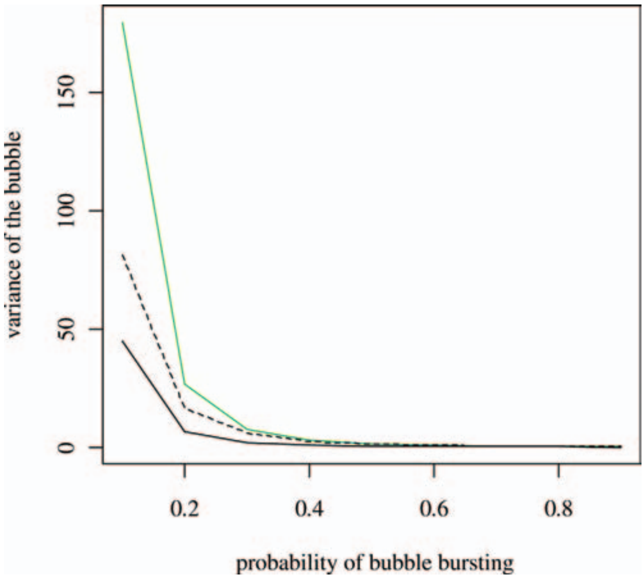
$$var(a_t) = \frac{q}{q+p}\left[\frac{\gamma^2(2-p)-\gamma^2}{p^2}\right]. \tag{25}$$

The variance of a_t is dependent on the probability of the bubble bursting, p, in a nonlinear way. An increase in p will reduce the variance of a_t and vice versa. In some ranges for p, a relatively small change in p will have a large effect on the variance of a_t. An increase in γ leads to a proportional shift up in the variance at every p. In figure 2 the solid black line is drawn for $q = 0.1$ and $\gamma = 1$. An increase in γ to 2 shifts the variance up to the grey line, where the variance at every p is $\gamma^2 = 4$ times larger than when $\gamma = 1$.

In contrast to the increase in γ, an increase in the probability of the bubble forming, q, has less effect on the variance of a_t. In figure 2, a large increase in q from 0.1 to 0.9, holding γ equal to 1, shifts the variance curve up from the solid line to the dashed line. The increase is proportionally larger the lower p is. For example, at $p = 0.9$ and $\gamma = 1$, the increase in q from 0.1 to 0.9 increases the relatively small variance by a factor of 5, while if $p = 0.1$, the variance only increases by a factor of 1.8.

Any change in a_t will be reflected in a change in ap_t; however, the bank does not know it is a_t that is changing. The bank chooses a volatility threshold beyond which it considers the movement in ap_t to be nonfundamental. This threshold is a number of standard deviations of ap_t. Up until this threshold, the bank ignores the changes

Figure 2. The Variance of the Bubble, a_t, versus the Probability of the Bubble Bursting, p



Note: $q = 0.1, \gamma = 1$ (solid black line); $q = 0.9, \gamma = 1$ (dashed line); $q = 0.1, \gamma = 2$ (grey line).

in ap_t, considering them most likely to be changes in fundamental values or noise. Beyond this threshold, the asset-price movement is regarded as nonfundamental. In this case, the bank takes account of ap_t in its estimates of y_t, i.e., in yf_t and in its optimal interest rate rule (i.e., $\kappa_t = 1$).

In choosing a threshold number of standard deviations, there is a trade-off between catching the bubble early and mistaking a fundamental asset-price movement for a nonfundamental one and acting when it is not necessary to do so. If a relatively low threshold is adopted, then the central bank will be more likely to catch bubbles early—before they send the economy a long way from equilibrium—but it is also more likely to make mistakes about asset-price bubbles. If the threshold is high, the bank will not catch the bubble as early but is less likely to make a mistake confusing a nonfundamental and fundamental price movement.

Given this trade-off, what might the appropriate threshold and its determinants be for the central bank to use when determining whether an asset-price movement is most likely a bubble or not? In the next section, simulation exercises are conducted to investigate these issues.

3. Simulation: Monetary Policy under Uncertainty about Asset-Price Shocks

In this section the model described above is simulated to determine what threshold will be appropriate and how sensitive the threshold choice is to the asset-price bubble parameters. The initial values for the model are $a_0^* = v_0 = y_0 = \pi_0 = a_0 = 0$; i.e., the economy is equilibrium, the output gap is 0, inflation is equal to target, and the market asset price is equal to its long-run average. Inflation and output have equal weight in the loss function; i.e., $\mu = 1$.

The model is simulated across different thresholds with varying parameter values for p, the probability of the bubble bursting, and γ, the amount the bubble grows each period. These two parameters are varied because, relatively, they have the most influence on the variance of a_t and, through this, the optimal threshold. Other parameters that affect the variance of ap_t—including q, the probability of a bubble forming, and τ, the autoregressive coefficient in the short-run fundamental asset-price equation—are fixed.[8] Changing q and τ, which is examined in the extensions subsection below, does not materially alter the main conclusions.

The model is simulated for 11,000 periods, and the first 1,000 values are discarded to remove any influence of initial values. The first step is to generate shocks to a_t^*, y_t, and π_t of 11,000 periods in length. The values of γ and p are fixed, and the model is simulated assuming the central bank cannot observe bubbles and uses monetary policy rules, which vary by the threshold number (one, two, three) of standard deviations of ap_t that the asset price must move before the central bank will consider it a bubble.

In reality, the central bank only has a limited amount of asset-price data to assess the variance of ap_t, so in the simulation the

[8]The assumed parameters are $q = 0.1$ and $\tau = 0.9$.

bank only uses the previous 100 years of ap_t to calculate the variance. This introduces some error into the estimation of $var(ap_t)$. Simulations where the central bank ignores asset-price movements completely in setting interest rates, and a base simulation where the central bank can identify bubbles, are also conducted. The different monetary policy rules are compared by calculating the welfare loss as measured by the bank's objective function (11). The values of γ and p are then changed, and the model is simulated again with the varying thresholds.

The welfare losses from using each monetary policy rule are shown across a variety of parameter values in table 1 and in the following figures.[9] Table 1 expresses the welfare losses as a ratio of the loss from following a particular rule to the loss if the central bank could identify bubbles and reacted optimally to them.

Comparing the three threshold-based monetary policy rules, the results in figure 3 show the threshold number of standard deviations of ap_t that will deliver the minimum welfare loss across a range of parameter values for the bubble. The parameter values vary from 1 to 8 for γ and 0.1 to 0.9 for p.[10] As can be seen from the plot, the threshold that delivers the minimum welfare loss depends on both p and γ. When p is high and γ is low—for example, $p = 0.9$

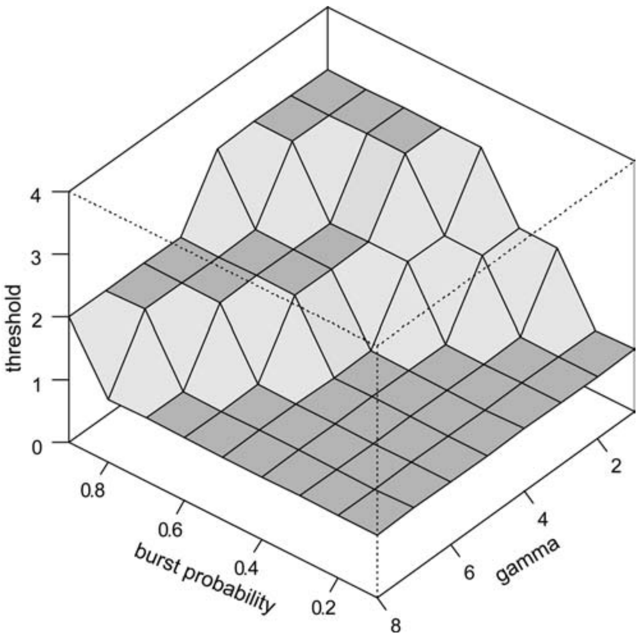
[9] A full set of results across the entire parameter range considered in the simulations is provided in appendix 3.

[10] With this parameter range, the expected effect of bubbles on the output gap ranges from 0.03 percent ($p = 0.9, \gamma = 1$) to 2.4 percent of GDP ($p = 0.1, \gamma = 8$), although actual bubble effects could be much larger in some cases. The larger expected shocks are around the maximum size of the U.S. output gap over the 1987 to 2007 period, when the output gap moved in a range of around ± 2 percent (OECD Economic Outlook database estimates). More importantly, the parameter range chosen for p and γ allows the testing of the various rules across a very wide range for the variance of the bubble relative to the total asset-price movement, as it is this relative variance that is the key determinant of the appropriate standard-deviation threshold for considering asset-price movements. In the simulations, the asset-price bubble variance as a percentage of the total asset-price variance ranges from less than 1 percent of the total asset-price variance ($p = 0.9, \gamma = 1$) to over 99 percent ($p = 0.1, \gamma = 8$). The results in the paper are robust to changing the range of output-gap effects that may be of interest to a central bank. For example, if ι (the marginal propensity to consume out of wealth) was increased, this would increase the size of the expected effect of the asset-price bubbles on the output gap at each γ and p, while not changing the overall results, as the relative variance of the bubble to the total asset price at each γ and p combination would remain the same.

Table 1. Welfare-Loss Ratio (Threshold/Optimal)

p	One Standard Deviation	Two Standard Deviations	Three Standard Deviations	Ignore Asset Price
$\gamma = 1$				
0.1	4.58	11.85	40.80	50.20
0.2	2.02	3.03	5.39	10.43
0.3	1.81	1.62	1.90	2.50
0.4	1.78	1.36	1.43	1.57
0.5	1.80	1.29	1.18	1.20
0.6	1.79	1.28	1.16	1.13
0.7	1.78	1.23	1.10	1.07
0.8	1.77	1.22	1.08	1.05
0.9	1.78	1.23	1.05	1.03
Mean	2.12	2.68	6.12	7.80
$\gamma = 3$				
0.1	13.91	70.18	194.38	238.16
0.2	3.01	9.78	23.88	46.68
0.3	1.73	3.32	5.68	15.00
0.4	1.53	1.79	2.46	5.91
0.5	1.52	1.38	1.76	3.07
0.6	1.64	1.34	1.50	1.98
0.7	1.66	1.25	1.28	1.46
0.8	1.67	1.23	1.21	1.38
0.9	1.73	1.23	1.15	1.16
Mean	3.15	10.17	25.92	34.98
$\gamma = 5$				
0.1	25.87	113.37	356.07	405.26
0.2	4.47	15.64	42.17	101.72
0.3	2.09	5.96	14.49	36.86
0.4	1.58	2.62	4.74	13.60
0.5	1.43	1.75	2.59	7.09
0.6	1.49	1.40	1.80	3.41
0.7	1.53	1.30	1.57	2.42
0.8	1.54	1.22	1.33	1.75
0.9	1.60	1.24	1.30	1.45
Mean	4.62	16.06	47.34	63.73

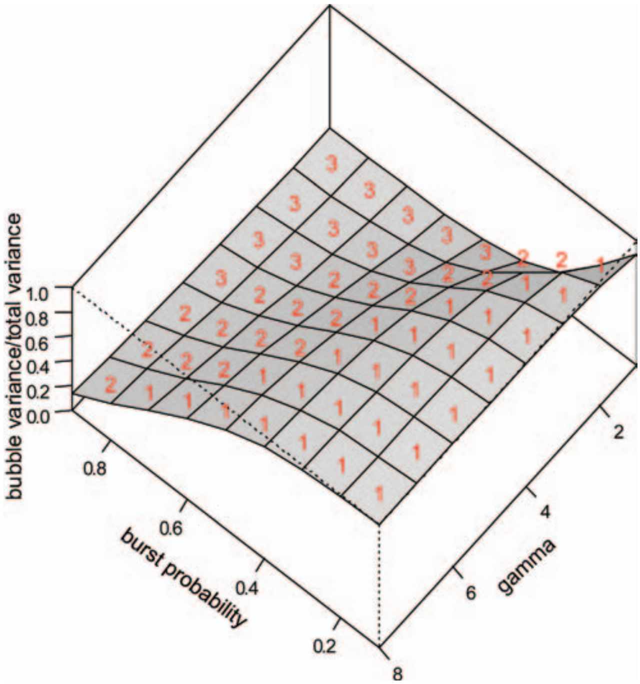
Figure 3. Thresholds Generating the Minimum Welfare Loss across the Parameters p and γ



and $\gamma = 1$—the central bank minimizes welfare losses with a conservative rule, only regarding an asset-price movement as a bubble if the asset price moves more than three standard deviations away from the long-run asset price. As p decreases and γ increases, the welfare-loss-minimizing threshold falls.

The parameters, p and γ, are important determinants of the threshold because they are key influences on the relative variances of the fundamental asset price and the bubble. The parameter γ has a positive nonlinear relationship with the variance of the bubble, while (as shown in figure 2) p has a negative nonlinear relationship with the bubble variance. The surface in figure 4 shows how the ratio of the bubble variance to the total asset-price variance depends on the parameters p and γ. When p is low and γ is high, the bubble has a high variance relative to the total asset-price variance. As p increases and γ decreases, the variance of the bubble relative to the total asset price decreases. Superimposed on the surface of figure 4

Figure 4. Relative Variance of the Bubble (Bubble Variance/Total Asset-Price Variance)



Note: Optimal thresholds are superimposed on the surface.

is the welfare-loss-minimizing threshold number of standard deviations. As the variance of the bubble relative to the total asset-price variance falls, the optimal threshold rises.

The relative variance of the bubble to the total asset price is an important determinant of the threshold because it determines the ease of distinguishing between fundamental and nonfundamental asset-price movements. For example, holding γ constant, increasing p decreases the variance of the bubble relative to the total asset-price variance, increasing the likelihood that any given asset-price movement is fundamental. In these circumstances, the bank should be more cautious about deciding that a given asset-price movement is nonfundamental and adopt a higher threshold number of standard deviations.

The dependence of the optimal threshold number of standard deviations on the probability of the bubble bursting, p, and the growth in the size of the bubble, γ, means that once a central bank has chosen or estimated the parameters of the stochastic process for a_t, it has effectively chosen its threshold for determining whether it should consider an asset-price movement fundamental or not.

In practice, it may be hard for a central bank to know with any accuracy what the value of the parameters p and γ are, so it needs to choose a monetary policy that will be robust to this parameter variation. The following discussion compares the different monetary policy rules across a range of parameter values for p and γ. Ignoring asset prices (i.e., giving them no separate weight in the monetary policy rule) is used as a benchmark because, as discussed in the paper, this is regarded by many in the literature as the optimal response to asset-price movements given the uncertainty about the nature of asset-price movements.

Figure 5 shows the ratio of the welfare loss from using a three-standard-deviation threshold compared with ignoring asset-price movements. Across a wide range of p and γ values, the welfare loss from using a three-standard-deviation rule is significantly less than the loss from ignoring asset prices. In some ranges the loss from the three-standard-deviation rule is only 30–40 percent of the loss incurred by ignoring asset prices. The four dots mark the area where the welfare loss is higher using the three-standard-deviation rule, but even in these cases the loss is only around 2 percent greater than the loss from ignoring asset prices. Furthermore, the absolute welfare gains to be made when the three-standard-deviation rule is superior to ignoring asset prices are far larger than the absolute losses when it is inferior. On average, the absolute welfare gain for the parameter combinations where the three-standard-deviation rule is superior to ignoring asset prices is more than 1,500 times the extra welfare losses for parameter combinations where it is inferior.[11] Over the whole parameter range, the total absolute loss from using the three-standard-deviation rule is 79 percent of the loss from ignoring asset prices.

[11] Welfare-loss averages referred to in the discussion of results throughout the paper are simple averages rather than weighted averages.

Figure 5. Ratio of the Welfare Loss from the Three-Standard-Deviation Rule to the Loss from Ignoring Asset Prices (3sd/ignore)

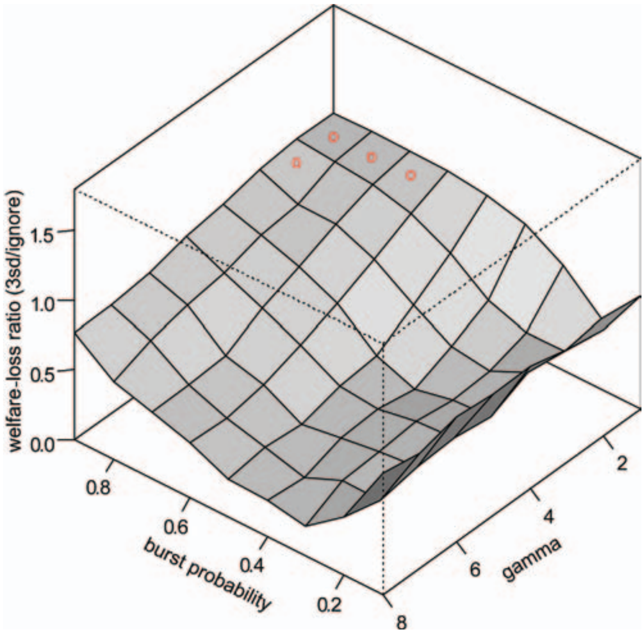
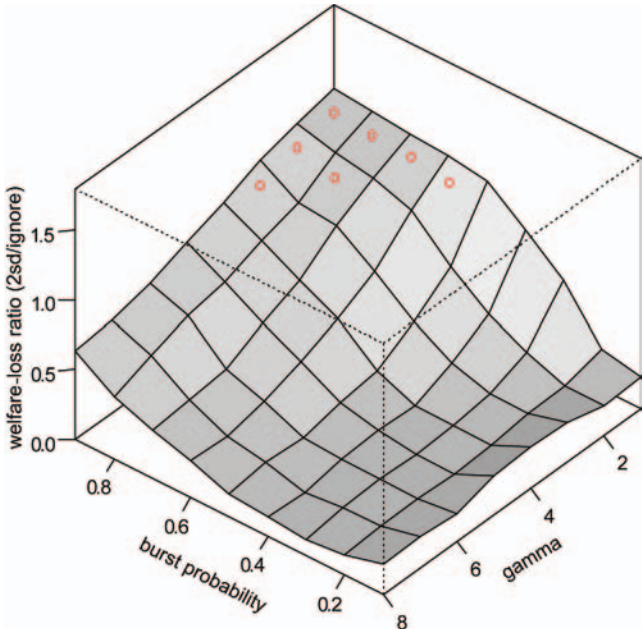


Table 1 and figures 6 and 7 show that as the threshold is lowered further, welfare losses relative to ignoring asset prices will decrease where the threshold rule is superior to ignoring asset prices, but there will be a greater range of bubble parameters where the loss from using the threshold rule will be higher. Figure 6 shows the ratio of welfare losses from using a two-standard-deviation rule. Over the whole parameter range, the total absolute loss from using the two-standard-deviation rule is 23 percent of the loss from ignoring asset prices altogether. However, compared with the three-standard-deviation rule, there is a greater parameter range (marked by the dots in figure 6) over which the welfare loss is greater for the threshold rule than it is for ignoring asset prices. As can be seen in figure 7, reducing the threshold to one standard deviation will further increase the parameter range over which using the threshold will lead to greater losses than ignoring asset prices (marked by the dots

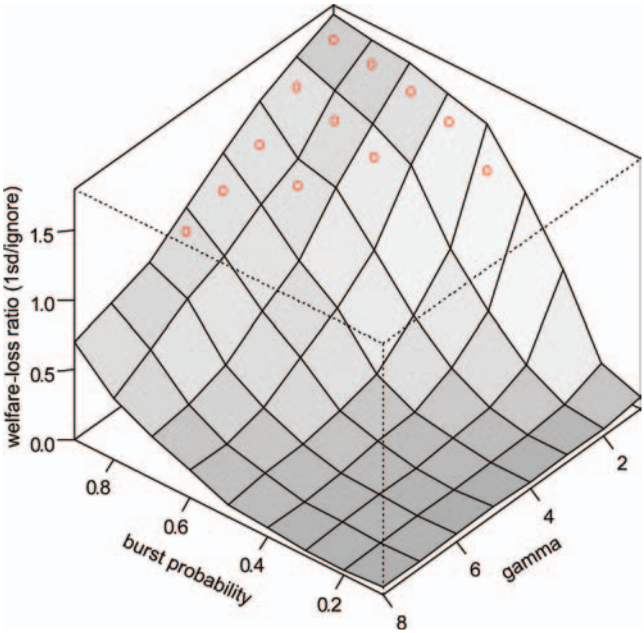
Figure 6. Ratio of the Welfare Loss from the Two-Standard-Deviation Rule to the Loss from Ignoring Asset Prices (2sd/ignore)



in figure 7). The size of these relative losses will also increase. Nevertheless, across the whole parameter range, the total absolute welfare loss will only be 6 percent of the loss from ignoring asset prices.

Overall, a three-standard-deviation rule performs better than ignoring asset prices across almost the entire parameter range for p and γ. Lowering the threshold below three standard deviations will decrease the total welfare loss further but increase the parameter range over which the loss from the threshold rule will be higher than the loss from ignoring asset prices. The results show that once the central bank has chosen or estimated the parameters of the asset-price process, it has effectively chosen its optimal threshold. If the central bank has little confidence in its estimates of the bubble parameters and simply assumes they are all equally likely, the threshold rule it chooses will depend on its degree of risk aversion. The more risk averse the central bank is, the higher the threshold it should use. A conservative approach would be to adopt a

Figure 7. Ratio of the Welfare Loss from the One-Standard-Deviation Rule to the Loss from Ignoring Asset Prices (1sd/ignore)



three-standard-deviation rule, as it will almost always do better than ignoring asset prices and result in only small absolute losses relative to ignoring asset prices in a small parameter range.

3.1 Further Extensions

This subsection examines alterations to a number of assumptions made above that affect the variance of the bubble, a_t, relative to the variance of ap_t, the total deviation in the market asset price from the long-run average, and therefore potentially the appropriate threshold. These include altering q, the probability of a bubble forming; τ, the autoregressive coefficient in the AR(1) equation; and a_t^* (the short-run deviation of the fundamental asset price) from the long-run average market asset price. Another simulation exercise is used to see the effect of changing the error distribution for a_t^*. In the original simulations, $q = 0.1, \tau = 0.9$, and the errors for a_t^* are assumed to be normally distributed. A simulation is also conducted to determine whether the threshold is sensitive to changing the bubble process to allow for asset-price overshooting and a_t falling below zero when the bubble bursts (i.e., a negative bubble).

3.1.1 Altering the Probability of a Bubble Forming, q

The model is simulated again over the range of 0.1 to 0.9 for p, $\gamma = 1$, and $q = 0.5$ and 0.9, i.e., with a 50 percent and 90 percent probability that a bubble will form every year when the economy is in the non-bubble state. All other parameters are the same as the original simulations. These results are compared with the original simulation with $\gamma = 1$ and $q = 0.1$. One of the new values for $q = 0.9$ is close to the top of its range, so this experiment gives the approximate limits of the effect of changing q. Table 2 shows the ratio of welfare losses for each parameter set. The effect of increasing q from 0.1 to 0.5 and 0.9 is to increase the range of p over which lower threshold rules of one and two standard deviations ensure the lowest welfare loss. This is because the increase in q increases the variance of a_t, the bubble, relative to $ap_t = a_t^* + a_t$, the deviation of the market asset price from the long-run average market asset price. This reduces the possibility of mistakenly deciding there is a bubble when there is not.

The general pattern of optimal thresholds remains the same as in the main simulations. In particular, with $q = 0.9$, the three-standard-deviation rule reduces welfare losses compared with ignoring asset prices across a wide range of p, with only small extra losses in a narrow p range. If $q = 0.9$, the range of p where there are greater losses from the three-standard-deviation rule is smaller than if $q = 0.1$, because the variance of the bubble is now larger. As in the main simulations, average welfare losses across the parameter range can be reduced further by decreasing the threshold below three, but this comes at the cost of a greater parameter range over which the threshold rule will lead to greater losses than incurred by ignoring asset prices.

3.1.2 Altering the Autoregressive Coefficient, τ

The model is simulated again across the range of 0.1 to 0.9 for p, $\gamma = 1$, and new lower autoregressive coefficients of $\tau = 0.1$

Table 2. Changing q: Welfare-Loss Ratio (Threshold/Optimal)

p	One Standard Deviation	Two Standard Deviations	Three Standard Deviations	Ignore Asset Price	Best Threshold Out of 1, 2, 3
$\gamma = 1$ and $q = 0.1$					
0.1	4.58	11.85	40.80	50.20	1
0.2	2.02	3.03	5.39	10.43	1
0.3	1.81	1.62	1.90	2.50	2
0.4	1.78	1.36	1.43	1.57	2
0.5	1.80	1.29	1.18	1.20	3
0.6	1.79	1.28	1.16	1.13	3
0.7	1.78	1.23	1.10	1.07	3
0.8	1.77	1.22	1.08	1.05	3
0.9	1.78	1.23	1.05	1.03	3
Mean	2.12	2.68	6.12	7.80	1
$\gamma = 1$ and $q = 0.5$					
0.1	10.39	34.17	90.40	93.30	1
0.2	3.25	7.68	14.79	18.80	1
0.3	2.33	3.20	4.58	5.84	1
0.4	2.01	2.14	2.66	3.18	1
0.5	1.94	1.76	1.88	2.00	2
0.6	1.87	1.54	1.58	1.61	2
0.7	1.85	1.40	1.34	1.35	3
0.8	1.81	1.33	1.22	1.21	3
0.9	1.80	1.29	1.18	1.15	3
Mean	3.03	6.06	13.29	14.27	1
$\gamma = 1$ and $q = 0.9$					
0.1	11.99	39.26	94.68	96.68	1
0.2	3.77	8.20	14.06	16.85	1
0.3	2.59	4.09	5.92	7.26	1
0.4	2.27	2.86	3.80	4.36	1
0.5	2.09	2.13	2.45	2.65	1
0.6	1.97	1.72	1.85	1.92	2
0.7	1.91	1.57	1.55	1.56	2
0.8	1.87	1.45	1.38	1.36	3
0.9	1.85	1.40	1.28	1.26	3
Mean	3.37	6.96	14.11	14.88	1

and 0.5 in the AR(1) that determines a_t^*. τ is set equal to 0.1 to find the approximate limits of the effect of changing this parameter. All other parameters are the same as in the original simulations. The new simulations are compared with the original simulation results, where τ was equal to 0.9. The results are in table 3.

Reducing τ decreases the variance of a_t^*, and so the variance of a_t rises as a proportion of the variance of $ap_t = a_t^* + a_t$. This reduces the probability of mistakenly identifying a given asset-price movement as a bubble when it is not, so lower thresholds produce the lowest welfare losses over a wider parameter range. At the extreme, if $\tau = 0.1$, it is relatively easy to identify bubbles, as most deviations from the long-run average asset price are bubble movements. A threshold of two standard deviations will provide a lower welfare loss than both ignoring asset prices and a three-standard-deviation rule across the entire range for p with $\gamma = 1$ and $\tau = 0.1$. The result that the three-standard-deviation rule is better than ignoring asset prices across the parameter range for p still remains, and it also has lower welfare losses than ignoring asset prices across the entire parameter range for p with $\tau = 0.1$. The advantage of the three-standard-deviation rule over the lower-standard-deviation rules is that variation in τ leads to less variation in its performance, compared with ignoring asset prices, than the other rules; i.e., it is less risky.

3.1.3 Changing the Error Distribution for a_t^*

In the main simulation, it is assumed that a_t^* is determined by an AR(1) with normally distributed errors. The model is simulated again with errors from a t distribution with eight degrees of freedom; i.e., the error distribution for the short-run deviation in the fundamental asset price has fatter tails than a normal distribution. This means that even in non-bubble periods, asset-price deviations from the long-run asset price have fat tails. The simulation is done with $\gamma = 1$ and over the range of 0.1 to 0.9 for p. All other parameter values are the same as in the original simulations. These results are compared with the original simulation in table 4.

The change in the error distribution from the normal to the t distribution raises the variance of a_t^* relative to ap_t. Overall, the

Table 3. Changing τ: Welfare-Loss Ratio (Threshold/Optimal)

p	One Standard Deviation	Two Standard Deviations	Three Standard Deviations	Ignore Asset Price	Best Threshold Out of 1, 2, 3
$\gamma = 1$ and $\tau = 0.9$					
0.1	4.58	11.85	40.80	50.20	1
0.2	2.02	3.03	5.39	10.43	1
0.3	1.81	1.62	1.90	2.50	2
0.4	1.78	1.36	1.43	1.57	2
0.5	1.80	1.29	1.18	1.20	3
0.6	1.79	1.28	1.16	1.13	3
0.7	1.78	1.23	1.10	1.07	3
0.8	1.77	1.22	1.08	1.05	3
0.9	1.78	1.23	1.05	1.03	3
Mean	2.12	2.68	6.12	7.80	1
$\gamma = 1$ and $\tau = 0.5$					
0.1	3.58	13.93	39.47	47.89	1
0.2	1.24	1.94	3.07	6.28	1
0.3	1.13	1.27	1.65	2.82	1
0.4	1.10	1.13	1.25	1.60	1
0.5	1.07	1.06	1.08	1.18	2
0.6	1.08	1.06	1.10	1.16	2
0.7	1.09	1.05	1.06	1.09	2
0.8	1.08	1.04	1.03	1.03	3
0.9	1.09	1.03	1.02	1.02	3
Mean	1.38	2.61	5.64	7.12	1
$\gamma = 1$ and $\tau = 0.1$					
0.1	3.13	11.52	44.64	56.54	1
0.2	1.20	1.91	3.60	7.77	1
0.3	1.08	1.26	1.58	2.87	1
0.4	1.05	1.12	1.23	1.63	1
0.5	1.05	1.07	1.12	1.31	1
0.6	1.06	1.04	1.07	1.13	2
0.7	1.05	1.03	1.05	1.07	2
0.8	1.05	1.02	1.03	1.04	2
0.9	1.05	1.02	1.02	1.02	2
Mean	1.30	2.33	6.26	8.26	1

Table 4. Changing Error Distributions for a_t^*

p	One Standard Deviation	Two Standard Deviations	Three Standard Deviations	Ignore Asset Price	Best Threshold Out of 1, 2, 3
$\gamma = 1$ and Normal Errors for a_t^*					
0.1	4.58	11.85	40.80	50.20	1
0.2	2.02	3.03	5.39	10.43	1
0.3	1.81	1.62	1.90	2.50	2
0.4	1.78	1.36	1.43	1.57	2
0.5	1.80	1.29	1.18	1.20	3
0.6	1.79	1.28	1.16	1.13	3
0.7	1.78	1.23	1.10	1.07	3
0.8	1.77	1.22	1.08	1.05	3
0.9	1.78	1.23	1.05	1.03	3
Mean	2.12	2.68	6.12	7.80	1
$\gamma = 1$ and t Distribution Errors for a_t^*					
0.1	4.06	12.27	33.00	42.77	1
0.2	2.23	2.35	3.94	7.33	1
0.3	2.23	1.70	2.01	2.62	2
0.4	2.23	1.60	1.69	2.01	2
0.5	2.23	1.39	1.25	1.27	3
0.6	2.21	1.29	1.13	1.13	3
0.7	2.20	1.29	1.09	1.06	3
0.8	2.20	1.28	1.07	1.04	3
0.9	2.22	1.28	1.06	1.02	3
Mean	2.42	2.71	5.14	6.69	1

lower-standard-deviation rules result in higher losses than previously because acting early is more likely to result in mistakenly treating an asset-price movement as a bubble when it is not, as it is now more likely the movement will actually be fundamental. The three-standard-deviation rule and ignoring asset prices now result in lower overall losses because it is more likely that the movements these rules fail to treat as a bubble are now fundamental. The pattern of three-standard-deviation rule losses compared with ignoring asset prices remains approximately the same as in the original simulations.

Relative to the original simulations, the above experiments show that a decrease in τ or an increase in q can increase the parameter range over which the one- and two-standard-deviation rules will result in the lowest welfare loss. However, changing the error distribution for a_t^* to a t distribution raised the welfare losses of these lower-standard-deviation rules relative to ignoring asset prices and using a three-standard-deviation rule. The main result, that the three-standard-deviation rule performs better than ignoring asset prices across most of the parameter range for variables that affect the variance of a_t relative to ap_t, remains once the parameter set is extended to q and τ.

3.1.4 Allowing for Asset-Price Overshooting and Negative Bubbles

Central banks assessing the effects of a potential bubble on the economy are often concerned about negative consequences for the real economy should the bubble burst. In the main simulation when the bubble bursts, the temporary positive boost to output from asset prices above fundamentals disappears and the output gap eventually returns to zero. For an economy that is used to much higher than equilibrium output, this is already likely to be regarded as a significant downturn. A more dramatic scenario, though, is that the asset price overshoots its fundamental value, leading to a negative bubble and output below equilibrium. To explore how this overshooting effect would affect welfare outcomes under the various thresholds, the model is simulated again with a bubble, a_t, that can become negative when the positive bubble bursts. In this simulation, if the bubble bursts in period t, in the following period a_{t+1} is given by

$$a_{t+1} = -b_t * a_t,$$

where b_t is a random uniform variable lying between 0 and 1, so that the negative bubble will be equal to between 0 and 100 percent of the positive bubble when the positive bubble bursts. If a negative bubble exists, it reduces by γ in the following periods until $a_t = 0$. While a negative bubble exists, $a_t < 0$, no positive bubble can form. Because the bubble can now become negative, this alters the expected effect of asset prices on the output gap next period and therefore the optimal interest rate rule, which for $a_t \geqslant 0$ becomes

$$r_t = \phi_1 y_t + \phi_2 \pi_t + \phi_3 (v_t \iota((1-p)(a_t+\gamma) - p b_t a_t) + (1-v_t)(\iota q \gamma)). \quad (26)$$

The main difference between the new rule and the original rule (15) is the term $-pb_ta_t$, which is added because when $v_t = 1$, with probability p_t, the bubble will burst and a_{t+1} will be equal to $-b_ta_t$ instead of 0 in the main simulation without a negative bubble. As in the GPS model, the expected effect on output of the bubble next period is the sum of the positive effect should the bubble continue and the negative effect should the bubble burst. The balance of these effects depends on the size of p and b_t. If the bubble has burst, while $a_t < 0$ interest rates need to offset the negative effect of the bubble on output next period, $\iota(a_t + \gamma)$, and the rule is

$$r_t = \phi_1 y_t + \phi_2 \pi_t + \phi_3(\iota(a_t + \gamma)). \tag{27}$$

If $ap_t \geqslant 0$, the activist central bank's interest rate rule is a modified version of (17):

$$rf_t = \phi_1 yf_t + \phi_2 \pi f_t + \kappa_t \phi_3 \iota((1 - pf_t)(ap_t + \gamma f_t) - pf_t b_t ap_t)), \tag{28}$$

where $\iota pf_t b_t ap_t$ is the expected effect of the bubble bursting on the output gap next period. The size of the negative bubble depends on the random variable b_t, which is set equal to its mean of 0.5. If there is a large downward shift in the asset price and $ap_t < 0$, the activist central bank will assume a bubble has burst and apply the following rule:

$$rf_t = \phi_1 yf_t + \phi_2 \pi f_t + \kappa_t \phi_3 \iota((ap_t + \gamma f_t)). \tag{29}$$

In this case the bank will treat the asset price as being in a negative bubble situation while the market asset-price deviation from the long-run asset price is more than a certain threshold number of standard deviations from its long-run average, and κ_t will be equal to 1 in this case. The results in table 5 show that the three-standard-deviation rule results in a lower welfare loss than ignoring asset prices across a very similar parameter range to the original simulations with no negative bubbles. With negative bubbles for $\gamma \leqslant 3$, p_t needs to be slightly smaller than without negative bubbles before the three-standard-deviation rule results in a lower welfare loss than ignoring asset prices; i.e., the variance of the bubble needs to be slightly larger relative to fundamental asset-price movements before a three-standard-deviation rule will be superior. Overall, the results of the simulation with the possibility of negative bubbles show that

Table 5. Simulation Allowing Negative Bubbles

p	One Standard Deviation	Two Standard Deviations	Three Standard Deviations	Ignore Asset Price	Best Threshold Out of 1, 2, 3
$\gamma = 1$					
0.1	7.48	24.91	53.28	57.81	1
0.2	2.48	2.86	4.01	5.47	1
0.3	2.50	1.67	1.82	2.27	2
0.4	2.62	1.50	1.34	1.35	3
0.5	2.66	1.44	1.17	1.15	3
0.6	2.69	1.40	1.10	1.09	3
0.7	2.70	1.38	1.07	1.04	3
0.8	2.68	1.39	1.05	1.01	3
0.9	2.69	1.39	1.05	1.01	3
Mean	3.17	4.22	7.74	8.02	1
$\gamma = 4$					
0.1	13.51	54.50	121.17	147.76	1
0.2	2.95	9.44	19.33	41.64	1
0.3	1.97	4.39	7.75	13.44	1
0.4	1.86	2.21	3.65	6.73	1
0.5	1.89	1.50	1.83	3.10	2
0.6	2.09	1.37	1.44	1.76	2
0.7	2.26	1.28	1.24	1.43	3
0.8	2.41	1.29	1.16	1.26	3
0.9	2.56	1.29	1.10	1.11	3
Mean	3.50	8.59	17.63	24.25	1
$\gamma = 8$					
0.1	18.51	72.06	212.81	246.32	1
0.2	4.95	24.81	50.46	85.82	1
0.3	1.88	5.42	11.19	27.10	1
0.4	1.64	2.97	5.56	13.94	1
0.5	1.46	1.90	3.00	5.93	1
0.6	1.51	1.55	2.09	4.06	1
0.7	1.61	1.39	1.65	3.07	2
0.8	1.79	1.27	1.31	1.69	3
0.9	1.98	1.25	1.21	1.35	3
Mean	3.93	12.51	32.14	43.25	1

the main result of the paper is also robust to a generalization of the asset-price process that allows for negative bubbles following a bubble bursting.

3.1.5 Other Possible Extensions

Dependence of Fundamental and Nonfundamental Asset-Price Movements. As stated in the paper, it is possible that fundamental and nonfundamental asset-price movements are related. A further extension of the model would be to allow dependence between fundamental and nonfundamental asset prices. This would introduce a covariance term in the variance of ap_t. If there was a positive covariance, then at least part of any asset-price movement would be fundamental and the variance of the bubble would be smaller relative to the total asset-price variance. In this case, the bank should be more cautious in deciding that an asset-price movement was partially nonfundamental and adopt a higher threshold number of standard deviations in its threshold rule. It would not alter the main conclusion that a three-standard-deviation rule is superior to ignoring asset prices. This is because while adding a covariance term would affect the ratio of the variance of the bubble to the total asset price, and therefore the optimal threshold rule, the three-standard-deviation rule has been shown to be robust across a wide range of values for the relative variance by altering other parameters, such as the bubble bursting.

Incorporating Interest Rate Effects on the Asset Price. It is unclear whether asset-price bubbles are affected by interest rates, and, in the model presented in this paper, it is assumed that the asset-price bubble is independent of interest rates. Although policymaker assumptions are often to the contrary, in reality, announced monetary policy and actual changes in interest rates often may not have any effect on investor behavior in a bubble situation. This can be examined from both the point of view of monetary policy expanding the bubble and monetary policy "popping" the bubble.

From the perspective of monetary policy inducing a larger bubble, in the 1990s it was argued that the knowledge that the Federal Reserve would cut interest rates in the event of a crash induced a bias toward gambling on continuing rises in the market.

However, this argument implies that monetary policy is effective very quickly in eliminating the effects of a crash, which market participants almost certainly know is not the case. Without a belief by the market that the central bank can eliminate the effects of a crash through interest rates, central bank announcements of future rate cutting in the event of a crash would not influence the bubble.

Furthermore, from the perspective of monetary policy directly shrinking or "popping" a bubble, if an asset-price bubble is in progress, investors will be earning and expecting large capital gains. These gains (often 20 or 30 percent or more per annum) are likely to more than cover any reasonable increase in interest rates, and expected gains will remain positive. In this situation, investors will continue to purchase the asset despite the monetary policy tightening, and so the bubble will continue.

However, it is also possible that changes in interest rates may indirectly affect asset prices in a bubble situation, particularly via their effect on aggregate demand and output growth. For example, the anticipation of a weaker-than-expected GDP growth outturn following monetary policy tightening may lead to a fall in investors' capital gains expectations.

A further extension of the model would be to incorporate interest rate effects on the asset price. Some potential ways to do this would be via interest rate effects on the probability of the bubble bursting or on the growth of the bubble. In this case, the required optimal interest rate changes would be less than currently produced by the model, as there would be two channels through which interest rates could affect aggregate demand (directly and indirectly via the asset price) and return the economy to potential. However, as above, incorporating interest rate effects on asset prices would not alter the main conclusion that a three-standard-deviation rule is superior to ignoring asset prices. This is because it has already been shown that the rule is robust to variation in the asset-price parameters that would most likely be affected by interest rates, such as the probability of the bubble bursting.

If the bubble is affected by interest rates, then there is no analytical solution such as that used in this paper that is available for the monetary policy rule, and numerical methods would be required to solve the model.

4. Conclusion

The results show that the threshold rule is sensitive to the assumptions about the asset-price change, particularly p, the probability of the bubble bursting, and γ, the growth of the bubble. To be consistent in its approach to monetary policy, the central bank should adjust not only its interest rate policy according to the assumptions it makes about the nature of the bubble process, as shown by Gruen, Plumb, and Stone (2005), but also its approach to determining when it will consider an asset-price movement most likely a bubble.

In practice, central banks face information constraints about the nature of asset-price shocks, and this complicates any active policy approach designed to take account of asset-price information. These results suggest that this problem should not lead to a policy approach where asset-price information is completely ignored prior to a bubble bursting.

A key finding is that it is better to adopt a high-threshold rule for deciding whether to include asset prices in the output-gap and inflation forecasts and monetary policy than to ignore asset prices altogether. Although it is difficult to determine the parameter assumptions for an asset-price bubble, a conservative three-standard-deviation threshold rule will result in lower welfare losses than ignoring asset prices altogether across a wide range of parameter values for the bubble.

Finally, the choice of threshold will depend on the central bank's degree of risk aversion. The more risk averse the bank, the higher the threshold it should adopt. Although a higher threshold will have a higher average absolute loss, it will have a smaller parameter range over which it will make a loss relative to ignoring asset prices.

Appendix 1. Solving for the Optimal Monetary Policy Rule

This appendix gives further detail on the dynamic programming method used to solve the linear-quadratic problem for the optimal monetary policy rule. The method described here closely follows Chow (1973). See also Bertsekas (2000). The economy is described by the system of linear equations, (9) and (10), expressed in matrix form below:

$$z_t = Az_{t-1} + Cx_t + b_t + \varepsilon_t, \tag{30}$$

where z_t is the vector of system variables, y_t and π_t, and x_t is a vector of control variables, r_t. b_t is a known series, in this case the expected effect of a_t on the output gap; ε_t is a vector of white-noise error terms; and A and C are coefficient matrices.

Written in matrix form, the objective function ((11) in the text) is to minimize

$$E_{t-1} \sum_{i=0}^{\infty} \delta^t z_t' K z_t, \tag{31}$$

where δ is the bank's discount factor and K is a diagonal matrix of weights on the system variables, in this case 1 and μ on y_t and π_t, respectively.

The overall problem is to find a rule for the control variables, x_t, that minimizes the objective function (31) subject to the system (30). The problem is solved backwards from the last period, T.

In period T the addition to the welfare function is $\delta^T z_T' K z_T$. Let $W_T = E_{T-1}(\delta^T z_T' K z_T)$. W_T is then written in the form

$$W_T = \delta^T E_{T-1}\big(z_T' H_T z_T + cst\big), \tag{32}$$

where $H_T = K$ and cst contains variables that do not involve x_T.

The term z_T in (32) is then replaced with the expression for z_T in (30) and expectations are taken. $E_{T-1}(\varepsilon_T) = 0$ by assumption, so this term is eliminated.

$$W_T = \delta^T [(Az_{T-1} + Cx_T + b_T)' H_T (Az_{T-1} + Cx_T + b_T) + cst] \tag{33}$$

To minimize W_T with respect to the control vector x_T, differentiate W_T and set it equal to 0:

$$\frac{\delta W_T}{\delta x_T} = 2C' H_T (Az_{T-1} + Cx_T + b_T) = 0. \tag{34}$$

Rearrange (34) to find the optimal policy at T, x_T:

$$\begin{aligned} C'HCx_T &= -C'H_TAz_{T-1} - C^{'}H_Tb_T \\ x_T &= -(C'H_TC)^{-1}C'H_TAz_{T-1} - (C'H_TC)^{-1}C'H_Tb_T \\ &= G_Tz_{t-1} + g_T, \end{aligned} \tag{35}$$

where

$$G_T = -(C'HC)^{-1}C'H_TA \tag{36}$$

$$g_T = -(C'HC)^{-1}C'H_Tb_T. \tag{37}$$

To find z_T under control, substitute the policy rule (35) into (30) to obtain

$$z_T = (A + CG_T)z_{t-1} + (Cg_T + b_T) + \varepsilon_T. \tag{38}$$

To find welfare, W_T, under control, substitute (38) into (32) to obtain

$$\begin{aligned} W_T = \delta^T E_{T-1}[((A + CG_T)z_{T-1} + (Cg_T + b_T) + \varepsilon_T)' \\ H_T((A + CG_T)z_{T-1} + (Cg_T + b_T) + \varepsilon_T] \end{aligned} \tag{39}$$

$$\begin{aligned} = \delta^T \big[((A + CG_T)z_{T-1})'H_T((A + CG_T)z_{T-1}) + \\ ((A + CG_T)z_{T-1})'H_T(Cg_T + b_T) + \\ (Cg_T + b_T)'H_T((A + CG_T)z_{T-1} + (Cg_T + b_T)' \\ H_T(Cg_T + b_T) + E_{T-1}\big(\varepsilon_T'\varepsilon_T\big)\big] \end{aligned} \tag{40}$$

$$\begin{aligned} = \delta^T[((A + CG_T)z_{T-1})'H_T((A + CG_T)z_{T-1}) \\ + z_{T-1}'(A + CG_T)'H_T(Cg_T + b_T) \\ + (Cg_T + b_T)'H_T(A + CG_T)z_{T-1}] + cst. \end{aligned} \tag{41}$$

$z_{T-1}'(A + CG_T)'H_T(Cg_T + b_T)$ is a scalar, so it is equal to its transpose:

$$\begin{aligned} W_T = \delta^T \big[\big(z_{T-1}'(A + CG_T)'\big)H_T((A + CG_T)z_{T-1}) \\ + 2z_{T-1}'(A + CG_T)'H_T(Cg_T + b_T)\big] + cst. \end{aligned} \tag{42}$$

Substitute (37) for g_T in (42):

$$W_T = \delta^T \left[\left(z'_{T-1}(A + CG_T)' \right) H_T((A + CG_T)z_{T-1}) \right.$$
$$\left. +2z'_{T-1}((A + CG_T)'H_T(C(-(C'H_TC)^{-1}C'H_Tb_T) + b_T) + cst. \right. \tag{43}$$

Rearrange $H_TC(C'H_TC)^{-1}C'$ as $C'H_TC(C'H_TC)^{-1} = I$, where I is the identity matrix.

$$W_T = \delta^T \left[\left(z'_{T-1}(A + CG_T)' \right) H_T((A + CG_T)z_{T-1}) + 2z'_{T-1} \right.$$
$$((A + CG_T)'((-C'H_TC(C'H_TC)^{-1}H_Tb_T + H_Tb_T)) + cst \tag{44}$$

$$= \delta^T \left[\left(z'_{T-1}(A + CG_T)' \right) H_T((A + CG_T)z_{T-1}) \right.$$
$$+ 2z'_{T-1}((A + CG_T)'(-H_Tb_T + H_Tb_T) + cst \tag{45}$$

$$= \delta^T \left[\left(z'_{T-1}(A + CG_T)' \right) H_T((A + CG_T)z_{T-1}) + cst, \right. \tag{46}$$

where (46) is the welfare loss at time T with x_T chosen optimally. The next step is to go back one period in time and choose x_{T-1}. We do not need to choose x_T, because it has already been chosen optimally no matter what the value of x_{T-1}. By Bellman's optimality principle we only need to choose x_{T-1} that will minimize the welfare loss at $T-1$ and T, as we know whatever value we choose for x_{T-1}, the value of x_T we have already determined will be optimal. The task is now to minimize W_{T-1} with respect to x_{T-1}:

$$W_{T-1} = \delta^{T-1} z'_{T-1} K z_{T-1} + \delta^T \left[\left(z'_{T-1}(A + CG_T)' \right) \right.$$
$$H_T((A + CG_T)z_{T-1}) + cst. \tag{47}$$

The first part of (47) is the welfare loss in $T-1$ and the second part is the welfare loss at T with x_T chosen optimally. We cannot improve welfare by changing x_T, but we can affect welfare in both T and $T-1$ with our choice of x_{T-1}.

Equation (47) can be rewritten as

$$W_{T-1} = \delta^{T-1} z'_{T-1} H_{T-1} z_{T-1} + cst, \tag{48}$$

where

$$H_{T-1} = K + \delta(A + CG_T)' H_T (A + CG_T)$$
$$= K + \delta A'[H_T - H_T C(C' H_T C)^{-1} C' H_T] A \qquad (49)$$
$$\text{(after substitution for } G_T\text{)}.$$

The minimization problem in (48) is of the same form as the original in (32), except now T has been replaced by $T-1$ so that x_{T-1} is therefore the same as for x_T, except T is now replaced by $T-1$ so that x_{T-1} is a function of G_{T-1}, which is in turn a function of H_{T-1}. The problem can be solved backwards in this iterative manner back to $T = 1$.

Equation (49) is the Ricatti equation. Bertsekas (2000) shows that this difference equation converges to the steady-state solution as $T \to \infty$:

$$H = K + \delta A'[H - HC(C'HC)^{-1}C'H]A \qquad (50)$$

and the optimal policy is

$$x_t = Gz_{T-1} + g \qquad (51)$$

$$G = -(CHC)^{-1}C'HA \qquad (52)$$

$$g = -(CHC)^{-1}C'Hb_t. \qquad (53)$$

This solution is implemented in Gauss by iterating (49) from initial values equal to the weights of y_t and π_t until it converges to a solution for H. Optimal policy is then calculated using (51)–(53).

Appendix 2. Calculating the Mean and Variance of the Bubble Process

This appendix describes how the mean and the variance of the stochastic bubble process, a_t, are derived.

The Asset-Price Bubble

The asset-price bubble, a_t, evolves according to a Markov chain, v_t:

$$\text{If } v_t = 1, \text{ then } a_t = a_{t-1} + \gamma. \qquad (54)$$

$$\text{If } v_t = 0, \text{ then } a_t = 0. \qquad (55)$$

That is, from an asset-price perspective there are two states of the world: bubble ($v_t = 1$) and no bubble ($v_t = 0$). The transition probability matrix between states is given by

	$v_{t+1} = 0$	$v_{t+1} = 1$
$v_t = 0$	$1 - q$	q
$v_t = 1$	p	$1 - p$

,

where p is the probability of the bubble bursting and q is the probability of the bubble forming.

The unconditional probability of each state is given:

$$pr(v_t = 1) = \frac{1 - (1 - q)}{2 - (1 - q) - (1 - p)}$$

$$pr(v_t = 1) = \frac{q}{q + p} \tag{56}$$

$$pr(v_t = 0) = \frac{p}{p + q}. \tag{57}$$

The Mean of a_t

The mean of the bubble process, $E(a_t)$, is given by summing across the states the unconditional probability of a state occurring, multiplied by the conditional mean of a_t given that state:

$$E(a_t) = pr(v_t = 1)E(a_t|v_t = 1) + pr(v_t = 1)E(a_t|v_t = 0). \tag{58}$$

Given $E(a_t|v_t = 0) = 0$, then

$$E(a_t) = pr(v_t = 1)E(a_t|v_t = 1), \tag{59}$$

where the unconditional probability of the bubble state, $pr(v_t = 1)$, is given above and the conditional mean of the bubble given the bubble state, $E(a_t|v_t = 1)$, is a geometric series given by

$$E(a_t|v_t = 1) = p\gamma \sum_{n=0}^{\infty}(1+n)(1-p)^n \text{ where } |1-p| < 1$$

$$E(a_t|v_t = 1) = \frac{p\gamma}{(1-(1-p))^2} \tag{60}$$

$$E(a_t|v_t = 1) = \frac{\gamma}{p}, \tag{61}$$

so the mean of a_t is given by

$$\begin{aligned} E(a_t) &= pr(v_t = 1)E(a_t|v_t = 1) \\ &= \frac{q\gamma}{(q+p)p}. \end{aligned} \tag{62}$$

The Variance of a_t

The variance of the bubble process, $var(a_t)$, is given by summing across the states the unconditional probability of a state occurring, multiplied by the conditional variance of a_t given that state:

$$\begin{aligned} var(a_t) = {} & pr(v_t = 1)\left[\left(E(a_t^2)\middle| v_t = 1\right) - (E(a_t|v_t = 1))^2\right] \\ & + pr(v_t = 0)\left[\left(E(a_t^2)\middle| v_t = 0\right) - (E(a_t|v_t = 0))^2\right], \end{aligned} \tag{63}$$

given $(E(a_t^2)|v_t = 0) = 0$ and $E(a_t|v_t = 0) = 0$, then

$$var(a_t) = pr(v_t = 1)\left[\left(E\left(a_t^2\right)\middle| v_t = 1\right) - (E(a_t|v_t = 1))^2\right], \tag{64}$$

where $(E(a_t^2)|v_t = 1)$ is a geometric series given by

$$\left(E\left(a_t^2\right)\middle| v_t = 1\right) = p\gamma^2 \sum_{n=0}^{\infty}(1+n)^2(1-p)^n \tag{65}$$

$$\left(E\left(a_t^2\right)\middle| v_t = 1\right) = p\gamma^2 \frac{1+(1-p)}{(1-(1-p))^3}$$

$$\left(E\left(a_t^2\right)\middle| v_t = 1\right) = \gamma^2 \frac{2-p}{p^2}, \tag{66}$$

so the variance of a_t is given by

$$var(a_t) = pr(v_t = 1)\left[\left(E\left(a_t^2\right)\middle| v_t = 1\right) - (E(a_t|v_t = 1))^2\right]$$
$$= \frac{q}{q+p}\left[\gamma^2 \frac{2-p}{p^2} - \left(\frac{\gamma}{p}\right)^2\right] \tag{67}$$
$$= \frac{q}{q+p}\left[\frac{\gamma^2(2-p) - \gamma^2}{p^2}\right]. \tag{68}$$

Appendix 3. Full Simulation Results

Table 6. Welfare-Loss Ratio (Threshold/Optimal)

p	**One Standard Deviation**	**Two Standard Deviations**	**Three Standard Deviations**	**Ignore Asset Price**
		$\gamma = 1$		
0.1	4.58	11.85	40.80	50.20
0.2	2.02	3.03	5.39	10.43
0.3	1.81	1.62	1.90	2.50
0.4	1.78	1.36	1.43	1.57
0.5	1.80	1.29	1.18	1.20
0.6	1.79	1.28	1.16	1.13
0.7	1.78	1.23	1.10	1.07
0.8	1.77	1.22	1.08	1.05
0.9	1.78	1.23	1.05	1.03
Mean	2.12	2.68	6.12	7.80
		$\gamma = 2$		
0.1	8.33	35.49	123.90	162.98
0.2	2.64	6.63	14.12	29.82
0.3	1.78	2.41	3.53	7.57
0.4	1.65	1.59	1.92	3.44
0.5	1.67	1.40	1.60	2.12
0.6	1.74	1.29	1.26	1.40
0.7	1.74	1.26	1.20	1.24
0.8	1.75	1.23	1.16	1.16
0.9	1.76	1.23	1.11	1.09
Mean	2.56	5.84	16.64	23.42

(continued)

Table 6. (Continued)

p	**One Standard Deviation**	**Two Standard Deviations**	**Three Standard Deviations**	**Ignore Asset Price**
		$\gamma = 3$		
0.1	13.91	70.18	194.38	238.16
0.2	3.01	9.78	23.88	46.68
0.3	1.73	3.32	5.68	15.00
0.4	1.53	1.79	2.46	5.91
0.5	1.52	1.38	1.76	3.07
0.6	1.64	1.34	1.50	1.98
0.7	1.66	1.25	1.28	1.46
0.8	1.67	1.23	1.21	1.38
0.9	1.73	1.23	1.15	1.16
Mean	3.15	10.17	25.92	34.98
		$\gamma = 4$		
0.1	21.42	141.66	420.60	474.87
0.2	3.64	12.49	27.65	72.81
0.3	1.73	3.62	7.05	19.79
0.4	1.59	2.20	3.63	12.39
0.5	1.46	1.57	2.14	5.40
0.6	1.59	1.37	1.66	2.63
0.7	1.62	1.33	1.52	2.01
0.8	1.64	1.24	1.30	1.53
0.9	1.68	1.22	1.20	1.29
Mean	4.04	18.52	51.86	65.86
		$\gamma = 5$		
0.1	25.87	113.37	356.07	405.26
0.2	4.47	15.64	42.17	101.72
0.3	2.09	5.96	14.49	36.86
0.4	1.58	2.62	4.74	13.60
0.5	1.43	1.75	2.59	7.09
0.6	1.49	1.40	1.80	3.41
0.7	1.53	1.30	1.57	2.42
0.8	1.54	1.22	1.33	1.75
0.9	1.60	1.24	1.30	1.45
Mean	4.62	16.06	47.34	63.73

(continued)

Table 6. (Continued)

p	One Standard Deviation	Two Standard Deviations	Three Standard Deviations	Ignore Asset Price
$\gamma = 6$				
0.1	28.22	103.35	452.13	505.21
0.2	5.14	19.13	44.13	111.88
0.3	2.43	7.39	16.02	43.00
0.4	1.59	3.19	6.26	19.13
0.5	1.44	1.99	3.00	10.30
0.6	1.41	1.58	2.19	5.13
0.7	1.44	1.34	1.67	3.31
0.8	1.50	1.26	1.41	1.95
0.9	1.54	1.27	1.40	1.72
Mean	4.97	15.61	58.69	77.96
$\gamma = 7$				
0.1	24.13	103.80	371.88	452.03
0.2	7.09	26.20	56.78	164.20
0.3	2.21	6.46	14.96	43.04
0.4	1.72	3.33	6.61	27.77
0.5	1.42	2.21	3.51	11.24
0.6	1.34	1.77	2.68	7.01
0.7	1.36	1.47	1.91	3.99
0.8	1.40	1.30	1.55	2.36
0.9	1.44	1.23	1.43	1.85
Mean	4.68	16.42	51.26	79.28
$\gamma = 8$				
0.1	25.40	112.87	460.97	524.93
0.2	6.38	24.33	68.51	162.77
0.3	2.68	7.05	13.67	63.11
0.4	1.81	4.20	7.84	29.69
0.5	1.52	2.56	4.45	16.30
0.6	1.35	1.76	2.70	6.79
0.7	1.35	1.55	2.15	4.55
0.8	1.33	1.39	1.73	2.10
0.9	1.42	1.29	1.56	2.03
Mean	4.81	17.44	62.62	90.36

References

Ball, L. 1999. "Efficient Rules for Monetary Policy." *International Finance* 2 (1): 63–83.

Bernanke, B., and M. Gertler. 1999. "Monetary Policy and Asset Price Volatility." *Economic Review* (Federal Reserve Bank of Kansas City) (4th Quarter): 17–51.

Bertsekas, D. P. 2000. *Dynamic Programming and Optimal Control.* Nashua, NH: Athena Scientific.

Cecchetti, S. G., H. Genberg, and S. Wadhwani. 2002. "Asset Prices in a Flexible Inflation Targeting Framework." NBER Working Paper No. 8970.

Chow, G. C. 1973. "Problems of Economic Policy from the Viewpoint of Optimal Control." *American Economic Review* 63 (5): 825–37.

Delong, J. B., A. Schleifer, L. H. Summers, and R. J. Waldmann. 1990. "Positive Feedback Investment Strategies and Destabilizing Rational Speculation." *Journal of Finance* 45 (2): 379–95.

Gruen, D., M. Plumb, and A. Stone. 2005. "How Should Monetary Policy Respond to Asset-Price Bubbles?" *International Journal of Central Banking* 1 (3): 1–32.

Poterba, J. M. 2000. "Stock Market Wealth and Consumption." *Journal of Economic Perspectives* 14 (2): 99–118.

Robinson, T., and A. Stone. 2005. "Monetary Policy, Asset-Price Bubbles and the Zero Lower Bound." Research Discussion Paper No. 2005-04, Reserve Bank of Australia.

Inflation: Do Expectations Trump the Gap?*

Jeremy M. Piger[a] and Robert H. Rasche[b]
[a]University of Oregon
[b]Federal Reserve Bank of St. Louis

We measure the relative contribution of the deviation of real activity from its equilibrium (the gap), "supply-shock" variables, and long-horizon inflation forecasts for explaining the U.S. inflation rate in the post-war period. For alternative specifications for the inflation-driving process and measures of inflation and the gap, we reach a similar conclusion: the contribution of changes in long-horizon inflation forecasts dominates that for the gap and supply-shock variables. Put another way, variation in long-horizon inflation forecasts explains the bulk of the movement in realized inflation. Further, we find evidence that long-horizon forecasts have become substantially less volatile over the sample period, suggesting that permanent shocks to the inflation rate have moderated. Finally, we use our preferred specification for the inflation-driving process to compute a history of model-based forecasts of the inflation rate. For both short and long horizons, these forecasts are close to inflation expectations obtained from surveys.

JEL Codes: C32, E31.

1. Introduction

The Phillips curve is one of the most recognized concepts in modern macroeconomics and is widely used as both a theoretical construct and an empirical tool. At the core of the Phillips curve is a relationship between inflation and the real activity "gap," defined as the

*We thank an anonymous referee for helpful suggestions. The views expressed here are those of the authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors. Piger: Department of Economics, 1285 University of Oregon, Eugene, OR 97403-1285 (jpiger@uoregon.edu); Rasche (corresponding author): Research Department, Federal Reserve Bank of St. Louis, P.O. Box 442, St. Louis, MO 63166 (rasche@stls.frb.org).

deviation of real economic activity from its equilibrium level. The within-sample statistical support for such a relationship in U.S. data over the post-war period is well documented in a number of studies, primary among them the work of Robert Gordon over the past twenty years (Gordon 1982, 1997, 1998). In particular, the gap is strongly statistically significant as an explanatory variable for inflation, and this significance is robust to a broad range of specifications of the Phillips curve. More recently, a number of papers have evaluated the out-of-sample forecasting performance of the Phillips curve. Here the evidence in favor of the gap as a driver for inflation is more mixed, with some papers documenting a substantial out-of-sample relationship (e.g., Stock and Watson 1999), while others find that inflation forecasts from a Phillips curve are not better than those from simple benchmark models such as a random walk or an autoregression (e.g., Atkeson and Ohanian 2001; Orphanides and van Norden 2005). Clark and McCracken (2006) provide a thorough exploration of the in-sample versus out-of-sample performance of the Phillips curve.

In this paper we revisit the importance of the gap as an explanatory variable for U.S. inflation over the post-war period. However, rather than measure importance with statistical significance, we instead focus on the relative contribution of the gap and other potential inflation drivers, such as changes in long-horizon inflation forecasts and "supply-shock" variables, for explaining the realized inflation rate. The initial analysis uses a specification for the inflation-driving process similar to that espoused by Gordon (1982, 1997, 1998). Subsequently, we investigate a specification that replaces the distributed lag on the inflation rate present in the Gordon specification with a time-varying intercept (TVI) that follows a random-walk process. The results from both the Gordon and TVI specifications are clear: changes in long-horizon inflation forecasts dominate the gap and supply-shock variables in the determination of actual inflation.[1]

[1]This result is reminiscent of findings in the bond-pricing literature that suggest that changes in long-horizon inflation expectations are the dominant source of variation in long-horizon bond yields (e.g., Gürkaynak, Sack, and Swanson 2005; Rudebusch and Wu 2008).

We then turn to more detailed analysis of the TVI model-based inflation forecasts. To begin, we allow for a sequence of structural breaks in the variance of shocks to the random-walk intercept. The estimates display a hump-shaped pattern, with the variance rising substantially during the late 1960s and the 1970s from its value in the 1950s and early 1960s, falling substantially in the early 1980s, and falling again in the early 1990s to its lowest level observed over the post-war period. This suggests that the size of permanent shocks with the inflation rate has varied substantially over the sample period.[2] Next, we use the TVI specification to construct histories of one-quarter-ahead inflation forecasts and compare these to survey-based inflation forecasts. For both short and long horizons, these forecasts are close to inflation expectations obtained from surveys, suggesting that the TVI model provides a reasonable description of the evolution of expectations.

The remainder of the paper proceeds as follows. Section 2 presents results for the Gordon-type Phillips-curve specification, while section 3 describes the TVI model and presents results from this specification. Section 4 compares the measures of inflation forecasts from the TVI model with survey-based measures of expected inflation. Section 5 concludes.

2. Results from the Gordon-Type Specification

2.1 Model Specification and Estimation

We begin with the specification that is featured in various analyses conducted by Robert Gordon:

$$\pi_t = a(L)\pi_{t-1} + b(L)D_t + c(L)X_t + \varepsilon_t. \tag{1}$$

Equation (1) relates the quarterly rate of inflation to a long (typically twenty-four quarters) distributed lag on inflation; an index of excess demand, D_t, measured as either the unemployment rate or the deviation of the unemployment rate from a time-varying NAIRU; and a vector of supply shocks, X_t, including changes in relative import

[2] Using a model with stochastic volatility, Stock and Watson (2007) also find substantial variability in the variance of shocks to the stochastic trend of inflation.

prices, changes in the relative price of food and energy, deviations of productivity from trend, and dummy variables for the beginning and termination of the Nixon price controls in the early 1970s. The distributed lag on inflation, $a(L)\pi_{t-1}$, is generally interpreted as "reflecting the influence of several past years of inflation behavior on current price-setting, through some combination of expectation formation and overlapping wage and price contracts" (Gordon 1998, 303).

Our specification differs from that in Gordon (1998) in that (i) it measures the gap using the "output gap," defined as the percentage deviation of real GDP from potential GDP as measured by the Congressional Budget Office (CBO); (ii) it uses four lags on all variables (in contrast to the twenty-four lags on inflation used by Gordon); and (iii) it does not include the productivity deviations present in the Gordon specification. To measure changes in relative import prices and the relative price of food and energy, we follow Gordon and use changes in import prices relative to the GDP price index and changes in the "core" PCE price index relative to the PCE price index. All the estimations follow Gordon and exclude a constant term.[3] We construct parallel analyses for the CPI, the PCE price index, and the GDP price index, each of which is measured in quarterly percentage changes at annual rates.

For presentation purposes, we focus on estimation of a transformed version of equation (1), which allows for direct estimation of the sum of the distributed lag coefficients, $a(1)$, $b(1)$, and $c(1)$:

$$\begin{aligned}\pi_t = {} & a(1)\pi_{t-1} + a^*(L)\Delta\pi_{t-1} + b(1)D_t \\ & + b^*(L)\Delta D_t + c(1)X_t + c^*(L)\Delta X_t + \varepsilon_t. \end{aligned} \tag{2}$$

Our estimates for equation (2) over the same 1962:Q1–1998:Q2 sample period used in Gordon (1998) are shown in table 1. The estimates of the sum of the distributed lag coefficients appear in bold.

In each of the three regressions, the sum of the estimated coefficients on lagged inflation is very close to unity, equaling 1.00 for

[3]Some initial regressions were constructed that included the constant term. The estimated constant was insignificant, and the estimates of the parameters of interest were unaffected by its omission.

Table 1. Gordon-Type Regressions
Sample Period: 1962:Q1–1998:Q2

	CPI	PCE	GDP
π_{t-1}	**1.01**	**1.00**	**1.00**
	(0.02)	(0.02)	(0.02)
$\Delta\pi_{t-1}$	−0.64	−0.67	−0.63
	(0.09)	(0.09)	(0.09)
$\Delta\pi_{t-2}$	−0.58	−0.44	−0.49
	(0.09)	(0.10)	(0.09)
$\Delta\pi_{t-3}$	−0.18	−0.27	−0.33
	(0.08)	(0.09)	(0.09)
Gap_t	**0.16**	**0.12**	**0.13**
	(0.04)	(0.03)	(0.04)
ΔGap_t	0.07	0.07	0.01
	(0.11)	(0.08)	(0.10)
ΔGap_{t-1}	0.13	0.01	0.01
	(0.11)	(0.08)	(0.10)
ΔGap_{t-2}	0.13	−0.01	0.08
	(0.11)	(0.08)	(0.10)
ΔGap_{t-3}	0.06	−0.12	0.08
	(0.11)	(0.08)	(0.10)
$\Delta Rel\ Import\ Prices_t$	**0.15**	**0.19**	**0.28**
	(0.11)	(0.09)	(0.11)
$\Delta^2 Rel\ Import\ Prices_t$	−0.07	−0.05	−0.42
	(0.11)	(0.08)	(0.10)
$\Delta^2 Rel\ Import\ Prices_{t-1}$	0.05	0.08	−0.17
	(0.09)	(0.07)	(0.09)
$\Delta^2 Rel\ Import\ Prices_{t-2}$	0.08	0.10	−0.05
	(0.08)	(0.06)	(0.08)
$\Delta^2 Rel\ Import\ Prices_{t-3}$	0.08	0.13	0.03
	(0.06)	(0.05)	(0.06)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_t$	**−0.15**	**0.20**	**−0.41**
	(0.90)	(0.71)	(0.86)
$\Delta^2 Rel\ Fd\ \&\ Energy\ Prices_t$	4.43	2.95	2.28
	(0.85)	(0.66)	(0.81)
$\Delta^2 Rel\ Fd\ \&\ Energy\ Prices_{t-1}$	3.36	2.08	1.68
	(0.88)	(0.68)	(0.77)

(*continued*)

Table 1. (Continued)

	CPI	PCE	GDP
Δ^2*Rel Fd & Energy Prices*$_{t-2}$	2.95	0.98	0.63
	(0.81)	(0.63)	(0.66)
Δ^2*Rel Fd & Energy Prices*$_{t-3}$	1.25	0.23	0.17
	(0.63)	(0.47)	(0.51)
NIXON_ON	−1.50	−1.19	−1.00
	(0.58)	(0.46)	(0.56)
NIXON_OFF	2.77	1.06	1.13
	(0.63)	(0.51)	(0.60)
$\overline{R}^2$	0.90	0.91	0.86
Std. Error of the Estimate	0.97	0.77	0.93
Durbin-Watson Statistic	2.13	2.06	2.14

Notes: This table shows OLS coefficient estimates and standard errors (in parentheses) for the Gordon Phillips-curve specification in equation (2) over the sample period considered in Gordon (1998), 1962:Q1–1998:Q2. Items in bold indicate estimates of the sum of the distributed lag coefficients for π_{t-1}, Gap_t, Δ*Rel Import Prices*$_t$, and Δ*Rel Fd & Energy Prices*$_t$. The "Gap" variable is measured using the estimate of the output gap produced by the CBO. All other variables are defined in section 2.

PCE and GDP inflation, and 1.01 for CPI inflation. The estimated sum of the coefficients on the output gap ranges from 0.12 to 0.16 and, consistent with prior research, is highly significant for all three price indices. The estimated sum of the coefficients on changes in relative import prices ranges from 0.15 to 0.28 and is significant in two of the three equations. The sign of the sum of the estimated coefficients on changes in the relative price of food and energy is not consistent across the three equations and is not significant in any equation, though the impact effect of this variable is always large and significant.[4]

[4]We assume that the output-gap and supply-shock variables are covariance stationary, and thus statements regarding statistical significance are based on standard Gaussian limiting distributions for t-statistics.

2.2 How Much Does the Gap Contribute to Explaining the Inflation Process?

We turn now to the relative contribution of the output gap for explaining inflation variability. To measure this relative contribution, we compute the "marginal adjusted R^2," defined as

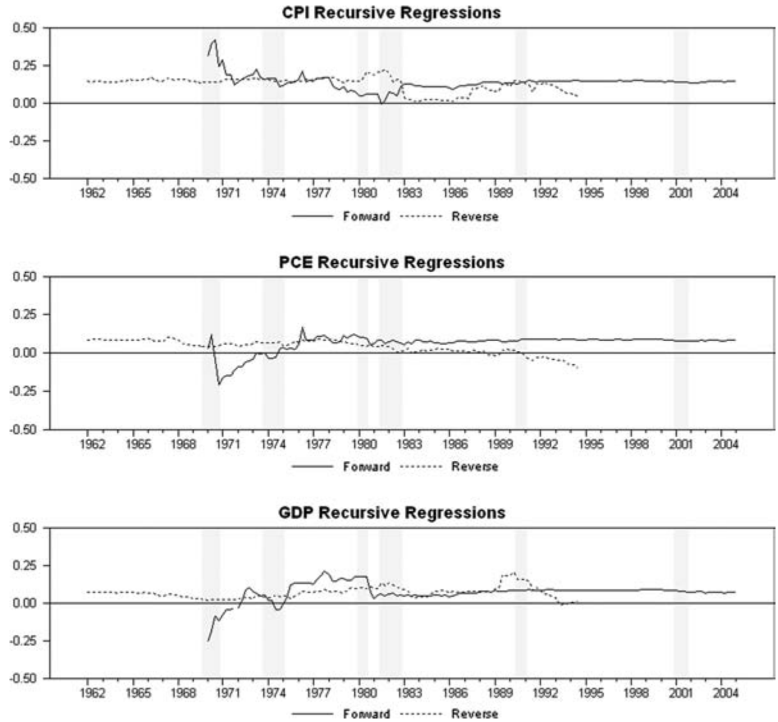
$$\overline{R}^2_m = \frac{\overline{R}^2_{\text{with gap}} - \overline{R}^2_{\text{without gap}}}{(1 - \overline{R}^2_{\text{without gap}})}, \tag{3}$$

where $\overline{R}^2_{\text{with gap}}$ is the adjusted R^2 from a regression with all the regressors including the distributed lag on the output gap, and $\overline{R}^2_{\text{without gap}}$ is the adjusted R^2 from a regression that excludes the distributed lag on the output gap. $\overline{R}^2_m$ gives the additional (adjusted) proportion of inflation variance explained by the model with output-gap terms included, measured relative to the (adjusted) amount of inflation variance left unexplained by the model that excludes output-gap terms. An $\overline{R}^2_m$ close to zero or negative indicates that the model that includes the output gap explains quantitatively little over the model that excludes the output gap, while an $\overline{R}^2_m$ close to one indicates that the addition of output-gap terms explains most of the inflation variance not explained by the model that excludes output-gap terms.

To investigate both full-sample and subsample contributions of the output gap, we compute $\overline{R}^2_m$ for forward and backward recursive regressions. In the forward recursions the sample period always begins in 1962:Q1. Initially, the sample ends in 1970:Q1 and then is extended one quarter at a time through 2005:Q1. In the backward recursive regressions, the sample size increases from the most recent observations. In all cases the end of the sample is fixed at 2005:Q1, and the beginning of the sample is initially 1994:Q3 and then shifted backward one quarter at a time until 1962:Q1. Figure 1 displays $\overline{R}^2_m$ for the forward and backward recursive regressions for each of the three measures of inflation.

Taken as a whole, the results in figure 1 suggest that the marginal explanatory power of the output gap is quantitatively small. Beginning with the forward regressions, for the PCE and GDP measures of inflation, $\overline{R}^2_m$ never exceeds 0.25 and is often even negative,

Figure 1. Marginal Adjusted R^2 Squares of Recursive Regressions



Notes: This figure shows recursive estimates of the marginal adjusted R^2 measure defined in equation (3) for the Gordon Phillips-curve specification given in equation (1). The solid line indicates forward recursive regressions beginning with the sample period 1962:Q1–1970:Q1 and ending with 1962:Q1–2005:Q1. The dotted line indicates reverse recursive regressions beginning with the sample period 1994:Q3–2005:Q1 and ending with 1962:Q1–2005:Q1.

indicating that the other regressors have a higher adjusted R^2 in the absence of the gap terms than does the full regression specification that includes the gap terms. For the longer sample regressions using the PCE or GDP measures of inflation, $\overline{R}_m^2$ is quite low, on the order of 0.08 to 0.10. For the CPI measure of inflation, the highest marginal contribution of the gap terms occurs for the shorter sample periods (late 1960s and 1970s), where at times $\overline{R}_m^2$ exceeds 0.30. For the longer samples, $\overline{R}_m^2$ is generally around 0.14, larger than that computed for the other two measures of inflation but still indicating relatively little marginal explanatory power for the output-gap terms.

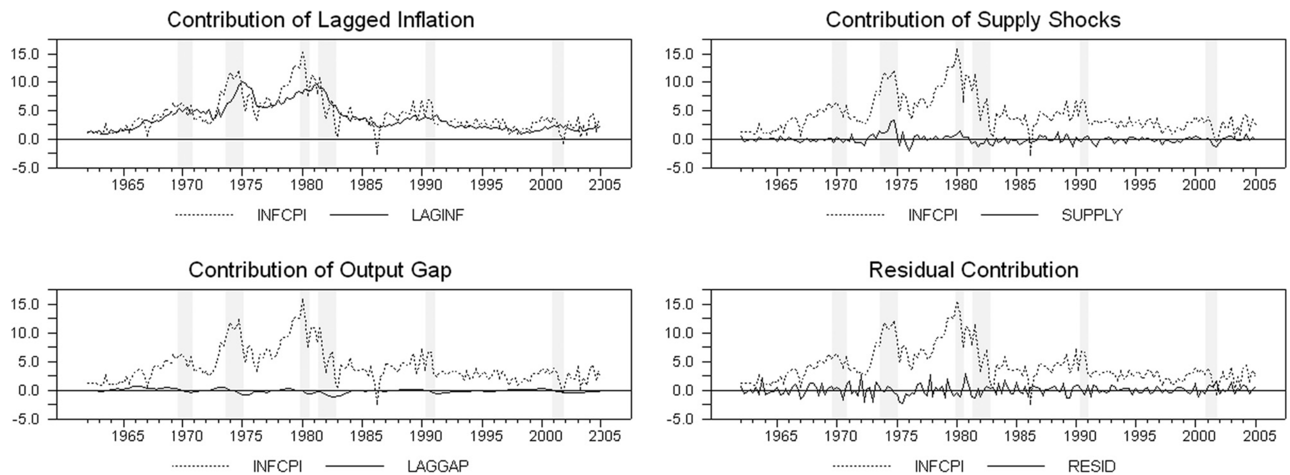
The marginal adjusted R^2 from the reverse recursive regressions does not alter the conclusions from the forward recursive regressions. Specifically, although $\overline{R}_m^2$ for the CPI and GDP measures of inflation is highly variable over recent sample periods, it never exceeds 0.25 and is usually much smaller. For the PCE measure of inflation, $\overline{R}_m^2$ is negative or close to zero for samples that include only recent years of data.

Another way to address this question is to compare the estimated values of $a(L)\pi_{t-1}$, $b(L)D_t$, $c(L)X_t$, and ε_t for a regression over the entire sample period. These are shown in figures 2–4 for regressions constructed on the sample 1962:Q1–2005:Q1. Note that for all three measures of inflation, the contribution of the gap terms, $b(L)D_t$, and the supply-shock variables, $c(L)X_t$, is dominated by the contribution of lagged inflation, $a(L)\pi_{t-1}$. These results are consistent with the analysis above: the output gap accounts for only a minor portion of fluctuations in inflation regardless of the measure of inflation. By contrast, inflation expectations, as proxied by a distributed lag on inflation whose coefficients sum to 1.0, account for the bulk of fluctuations in inflation. These results present a preliminary answer to the question posed in the title. Based on the Phillips-curve specification considered here, expectations do appear to trump the gap.

3. Results from the Time-Varying Intercept Specification

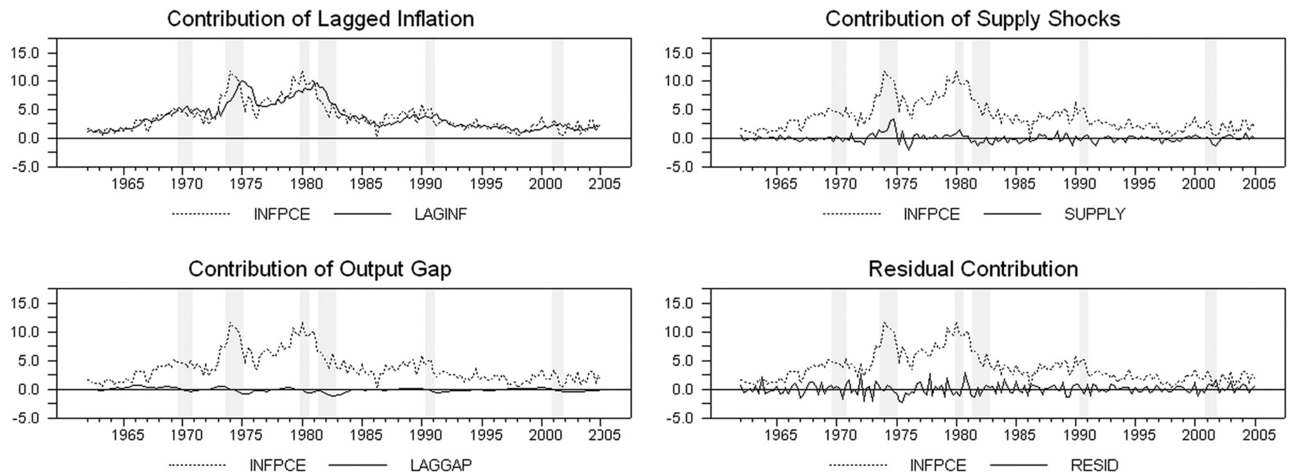
The results from the Phillips-curve specification in equation (1) indicate that inflation expectations are a dominant driver of realized inflation. In this section we refine this result using an alternative specification that allows us to focus more directly on the importance of movements in long-horizon inflation expectations. In particular, we extract a permanent random-walk component from the inflation process that can be interpreted as the long-horizon forecast of inflation. This allows us to directly assess the variability of changes in long-horizon inflation expectations as well as to investigate changes in this variability over time.

Figure 2. CPI Inflation: Gordon Equation, Sample Period 1962:Q1–2005:Q1



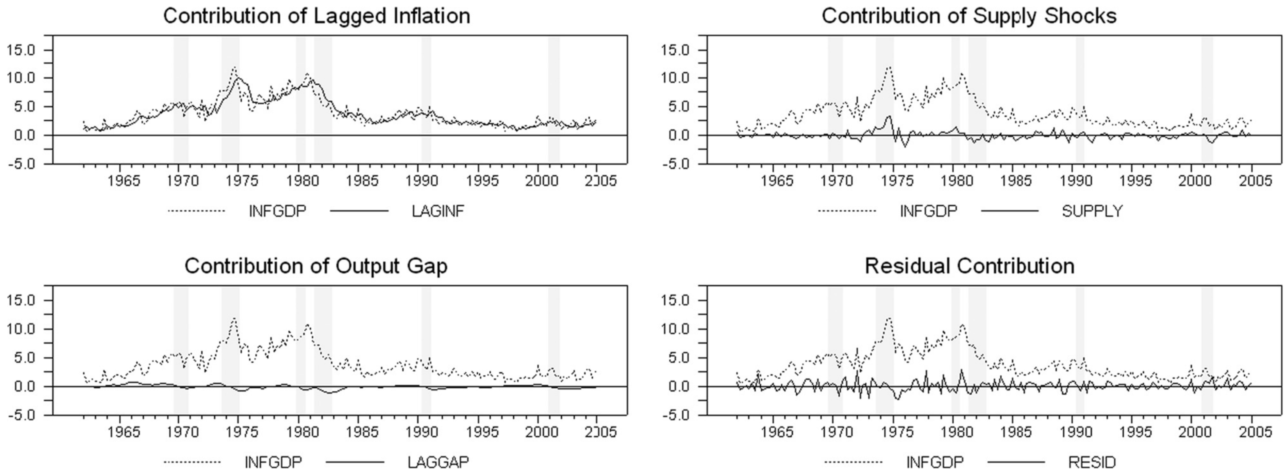
Notes: This figure plots the inflation rate (dotted line) against the estimated components of the Gordon Phillips-curve regression in equation (1), where parameter estimates were constructed using the sample period 1962:Q1–2005:Q1. The inflation measure is the CPI. The "Output Gap" variable is measured using the estimate of the output gap produced by the CBO. The "Supply Shock" variables are defined in section 2 above.

Figure 3. PCE Inflation: Gordon Equation, Sample Period 1962:Q1–2005:Q1



Notes: This figure plots the inflation rate (dotted line) against the estimated components of the Gordon Phillips-curve regression in equation (1), where parameter estimates were constructed using the sample period 1962:Q1–2005:Q1. The inflation measure is the PCE index. The "Output Gap" variable is measured using the estimate of the output gap produced by the CBO. The "Supply Shock" variables are defined in section 2 above.

Figure 4. GDP Inflation: Gordon Equation, Sample Period 1962:Q1–2005:Q1



Notes: This figure plots the inflation rate (dotted line) against the estimated components of the Gordon Phillips-curve regression in equation (1), where parameter estimates were constructed using the sample period 1962:Q1–2005:Q1. The inflation measure is the GDP index. The "Output Gap" variable is measured using the estimate of the output gap produced by the CBO. The "Supply Shock" variables are defined in section 2 above.

In particular, suppose that the distributed lag on inflation in the Gordon specification represents a proxy for long-horizon expected inflation that is specified to appear with a coefficient of 1.0 so that the long-run Phillips curve is vertical:

$$\pi_t = 1.0\pi_t^e + b(L)D_t + c(L)X_t + \varepsilon_t. \tag{4}$$

Alternatively, this equation can be thought of as specifying a time-varying intercept (TVI) on a vector of 1.0s:

$$\pi_t = 1.0z_t + b(L)D_t + c(L)X_t + \varepsilon_t. \tag{5}$$

We assume that z_t follows a random walk:[5]

$$z_t = z_{t-1} + \omega_t. \tag{6}$$

Equation (6) implies that, assuming stationarity of D_t and X_t, the infinite-horizon forecast of inflation is equal to z_t plus a constant term reflecting the unconditional mean of D_t and X_t (see Beveridge and Nelson 1981). Thus, variation in z_t has the interpretation of variation in the long-horizon inflation expectation.[6]

We estimate the model in equations (5)–(6) via maximum likelihood using the Kalman filter. The estimates of the model parameters are shown in table 2 for the sample period 1962:Q1–2005:Q1. Table 2 also shows the standard error of the estimate for the Gordon equation estimated over the same sample period, which demonstrates that the time-varying intercept specification is competitive with the Gordon specification.

---

[5]This is similar to Gordon's specification of the time-varying NAIRU in his 1997 and 1998 papers.

[6]Equation (6) assumes that the shocks to long-horizon inflation expectations are frequent and continuous. An alternative is that shocks to long-horizon inflation expectations are infrequent and discrete. For an example of such a specification for modeling U.S. inflation, see Levin and Piger (2002, 2005).

Table 2. Time-Varying Intercept Model
Sample Period: 1962:Q1–2005:Q1

	CPI	**PCE**	**GDP**
Standard Deviation of Intercept	0.58	0.37	0.35
	(0.08)	(0.05)	(0.05)
Gap_t	0.18	0.02	0.00
	(0.10)	(0.06)	(0.00)
Gap_{t-1}	0.00	0.00	0.01
	(0.00)	(0.00)	(0.07)
Gap_{t-2}	0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)
Gap_{t-3}	0.13	−0.12	0.07
	(0.12)	(0.08)	(0.00)
Gap_{t-4}	0.04	0.25	0.07
	(0.11)	(0.08)	(0.09)
$\Delta Rel\ Import\ Prices_t$	0.12	0.13	−0.20
	(0.06)	(0.04)	(0.04)
$\Delta Rel\ Import\ Prices_{t-1}$	0.07	0.09	0.06
	(0.06)	(0.04)	(0.05)
$\Delta Rel\ Import\ Prices_{t-2}$	0.07	0.02	0.07
	(0.07)	(0.05)	(0.05)
$\Delta Rel\ Import\ Prices_{t-3}$	0.03	0.05	0.11
	(0.06)	(0.04)	(0.05)
$\Delta Rel\ Import\ Prices_{t-4}$	0.00	0.04	0.05
	(0.05)	(0.04)	(0.04)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_t$	2.83	1.97	1.30
	(0.33)	(0.23)	(0.25)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-1}$	0.35	0.40	0.68
	(0.34)	(0.23)	(0.25)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-2}$	0.19	0.21	0.01
	(0.36)	(0.24)	(0.19)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-3}$	0.19	0.03	−0.03
	(0.35)	(0.23)	(0.24)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-4}$	0.08	0.24	0.29
	(0.27)	(0.22)	(0.25)
NIXON_ON	−0.84	−0.81	−1.27
	(0.97)	(0.65)	(0.67)
NIXON_OFF	3.03	2.10	2.54
	(0.79)	(0.50)	(0.59)

(*continued*)

Table 2. (Continued)

	CPI	**PCE**	**GDP**
Log-Likelihood	−276.16	−209.47	−225.91
Std. Error of the Estimate	0.90	0.64	0.78
Std. Error of the Estimate (Gordon Equation)	0.93	0.74	0.90

Notes: This table shows maximum likelihood coefficient estimates and standard errors (in parentheses) for the time-varying intercept Phillips-curve specification in equations (5)–(6) over the sample period 1962:Q1–2005:Q1. "Standard deviation of intercept" refers to the standard deviation of the innovation to the random-walk intercept term in equation (5). The "Gap" variable is measured using the estimate of the output gap produced by the CBO. All other variables are defined in section 2.

We focus our analysis on an expanded version of the TVI specification, the results of which are presented in table 3. First, we extend the sample period to include data subsequent to the end of the Korean War. Since the core PCE data are not available before 1959, we recompute the relative change in food and energy prices using CPI data. The "core CPI" is available starting in 1957. Prior to 1957 we use the "all items CPI less food" rather than the "core CPI." The two series are highly correlated in the late 1950s, since energy prices were not highly volatile until the early 1970s. Prior to 1987 we compute the relative change in food and energy prices using CPI data on a 1967 = 100 base, not seasonally adjusted, and apply the current seasonal factors for these years using the 1982–84 base-year data. We do this to avoid the truncation problems that affect the computation of CPI inflation rates in the early part of the sample period when the base year is 1982−84 = 100 (see Kozicki and Hoffman 2004).

Second, we allow for structural breaks in the variance of the innovations to the time-varying intercept process to occur at several points in the sample that align with well-known macroeconomic and monetary events. The first break is allowed to occur at the beginning of the Great Inflation, which we date to the first quarter of 1967. The second break is meant to capture the beginning of the large reduction in U.S. macroeconomic volatility that has been observed over the past two decades. Based on the findings of Kim and Nelson (1999)

Table 3. Time-Varying Intercept Model with Variance Breaks, Sample Period: 1953:Q1–2005:Q1

	CPI	PCE	GDP
Standard Deviation of Intercept 53–66	0.79	0.47	0.52
	(0.15)	(0.12)	(0.15)
Standard Deviation of Intercept 67–83	1.92	1.06	0.75
	(0.19)	(0.16)	(0.21)
Standard Deviation of Intercept 84–93	0.36	0.26	0.22
	(0.09)	(0.06)	(0.08)
Standard Deviation of Intercept 94–05	0.15	0.12	0.09
	(0.05)	(0.06)	(0.08)
Gap_t	−0.01	−0.09	−0.10
	(0.07)	(0.07)	(0.09)
Gap_{t-1}	0.10	0.22	0.13
	(0.09)	(0.09)	(0.10)
Gap_{t-2}	0.03	0.03	−0.03
	(0.10)	(0.09)	(0.11)
Gap_{t-3}	0.10	−0.11	0.03
	(0.10)	(0.09)	(0.12)
Gap_{t-4}	−0.07	0.06	0.01
	(0.08)	(0.07)	(0.09)
$\Delta Rel\ Import\ Prices_t$	0.10	0.07	−0.23
	(0.04)	(0.04)	(0.05)
$\Delta Rel\ Import\ Prices_{t-1}$	−0.002	0.08	0.11
	(0.04)	(0.04)	(0.05)
$\Delta Rel\ Import\ Prices_{t-2}$	0.04	−0.03	0.01
	(0.04)	(0.05)	(0.05)
$\Delta Rel\ Import\ Prices_{t-3}$	−0.06	0.04	0.09
	(0.04)	(0.04)	(0.05)
$\Delta Rel\ Import\ Prices_{t-4}$	0.03	−0.01	0.02
	(0.04)	(0.04)	(0.05)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_t$	3.26	2.14	1.36
	(0.21)	(0.18)	(0.22)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-1}$	0.21	0.01	0.17
	(0.21)	(0.13)	(0.23)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-2}$	−0.07	0.01	0.20
	(0.21)	(0.32)	(0.22)
$\Delta Rel\ Fd\ \&\ Energy\ Prices_{t-3}$	0.32	−0.21	0.03
	(0.21)	(0.19)	(0.33)

(*continued*)

Table 3. (Continued)

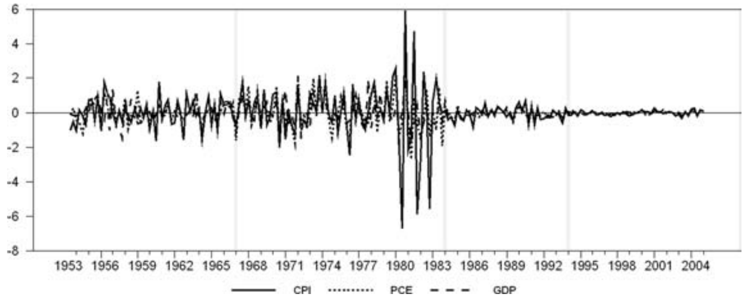
	CPI	PCE	GDP
Δ*Rel Fd & Energy Prices*$_{t-4}$	−0.18	0.30	0.41
	(0.21)	(0.18)	(0.22)
NIXON_ON	−0.36	−0.31	−1.51
	(1.75)	(1.08)	(0.96)
NIXON_OFF	3.37	2.20	2.46
	(1.16)	(0.73)	(0.71)
Log-Likelihood	−290.16	−249.08	−277.11
Std. Error of the Estimate	0.40	0.47	0.66

Notes: This table shows maximum likelihood coefficient estimates and standard errors (in parentheses) for the time-varying intercept Phillips-curve specification in equations (5)–(6) over the sample period 1953:Q1–2005:Q1, where the standard deviation of the innovation to the random-walk intercept term (denoted "Standard Deviation of Intercept") is allowed to change in 1967, 1984, and 1994. The "Gap" variable is measured using the estimate of the output gap produced by the CBO. All other variables are defined in section 2.

and McConnell and Pérez-Quirós (2000), we date the beginning of this "Great Moderation" to the first quarter of 1984. We date the third break at the first quarter of 1994, when the Federal Open Market Committee (FOMC) started releasing information on changes in the intended federal funds rate at the close of FOMC meetings.

As table 3 demonstrates, for all three measures of inflation the estimated variance of the innovations to the time-varying intercept increases sharply during the Great Inflation, falls to 40–50 percent of its 1953–66 value during the first decade of the Great Moderation, and then declines by roughly 50 percent of the value in the 1984–93 period during the most recent decade (see figure 5 for a plot of the estimated innovations). This pattern for the volatility of shocks to the random-walk intercept suggests that the size of permanent shocks to the inflation rate has varied substantially over the sample period, and that such shocks are now quite small from a historical perspective. The latest decline in volatility is consistent with

Figure 5. Shocks to Permanent Inflation



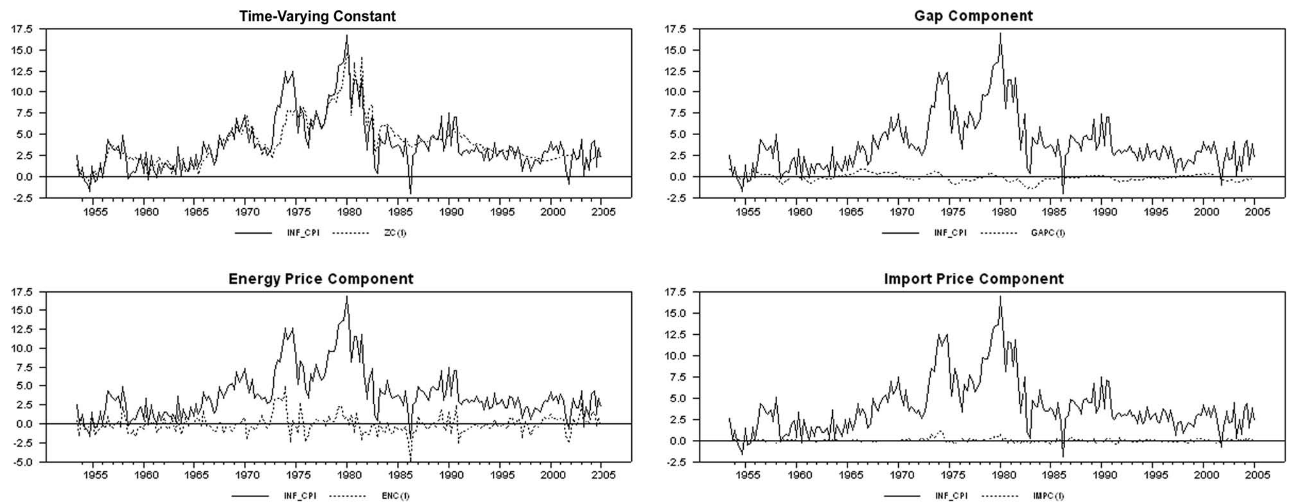
Notes: This figure plots the estimated innovations to the random-walk intercept of the time-varying intercept Phillips-curve specification in equations (5)–(6), where estimation is based on the sample period 1953:Q1–2005:Q1.

the notion that long-horizon inflation expectations have become better "anchored" during the period of increasing FOMC transparency, although this is not necessarily evidence of a causal relationship between increased transparency and lower volatility of long-term inflation expectations.

The estimates of the time-varying intercept and the contributions of the gap and supply shocks from the estimates in table 3 are shown in figures 6–8 for the three measures of inflation. These graphs indicate that the time-varying intercept term dominates the variation in all three measures of inflation. The only cases where the distributed lags on the output-gap and the supply-shock terms account for a substantial portion of the inflation rates are in 1973–74 and, to a lesser extent, in 1979–80.

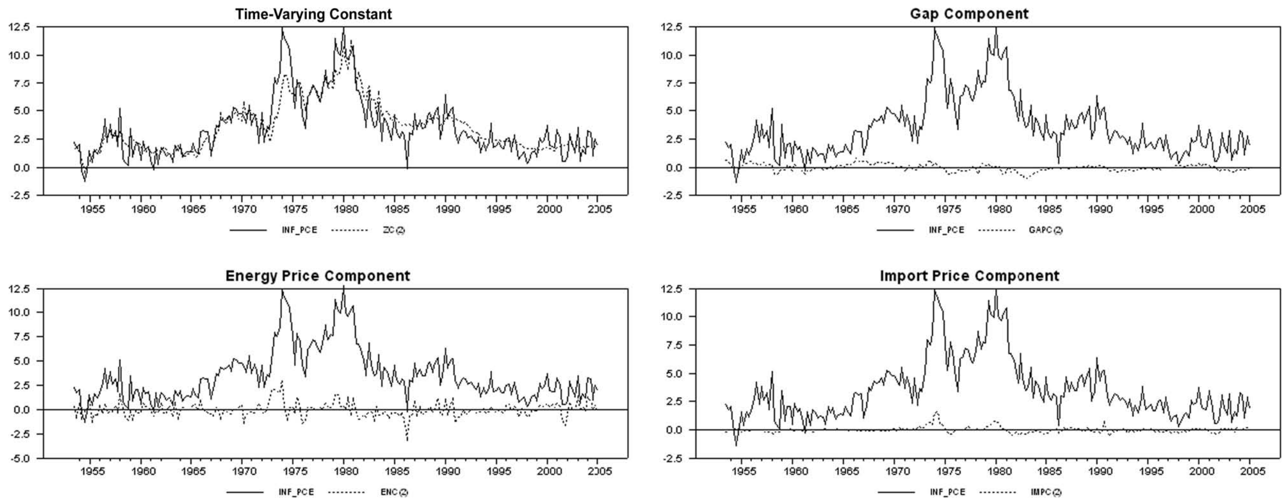
Finally, we have also estimated a version of the TVI specification in which the CBO measure of the output gap is replaced by the difference between the unemployment rate and a time-varying estimate of the NAIRU. We follow Gordon (1997) and model the NAIRU as a random walk and constrain the standard deviation of the error term in this process to 0.2. Results from this specification (not reported here) are substantially the same as those obtained with the CBO output gap, suggesting that our conclusions about the contribution of the gap are not sensitive to whether it is measured as an output or unemployment gap.

Figure 6. One-Period-Ahead CPI Inflation Components



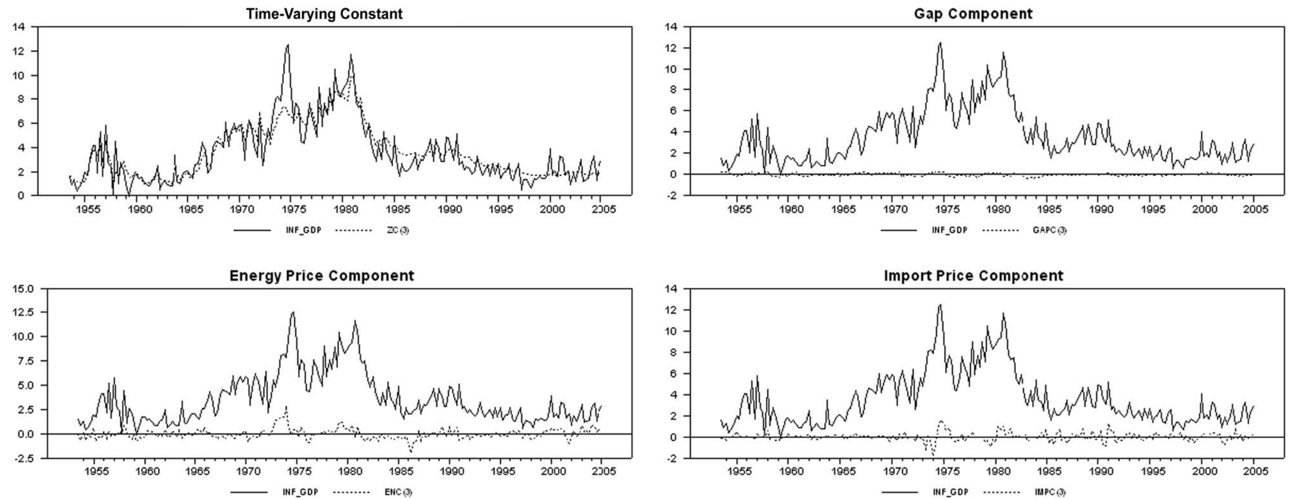
Notes: This figure plots the inflation rate (solid line) against the estimated components of the time-varying intercept Phillips-curve specification in equations (5)–(6), where parameter estimates were constructed using the sample period 1953:Q1–2005:Q1. The inflation measure is the CPI. The "Gap" variable is measured using the estimate of the output gap produced by the CBO. The "Energy Price" and "Import Price" variables are defined in section 2 above.

Figure 7. One-Period-Ahead PCE Inflation Components



Notes: This figure plots the inflation rate (solid line) against the estimated components of the time-varying intercept Phillips-curve specification in equations (5)–(6), where parameter estimates were constructed using the sample period 1953:Q1–2005:Q1. The inflation measure is the PCE index. The “Gap” variable is measured using the estimate of the output gap produced by the CBO. The “Energy Price” and “Import Price” variables are defined in section 2 above.

Figure 8. One-Period-Ahead GDP Inflation Components



Notes: This figure plots the inflation rate (solid line) against the estimated components of the time-varying intercept Phillips-curve specification in equations (5)–(6), where parameter estimates were constructed using the sample period 1953:Q1–2005:Q1. The inflation measure is the GDP index. The "Gap" variable is measured using the estimate of the output gap produced by the CBO. The "Energy Price" and "Import Price" variables are defined in section 2 above.

4. TVI Model-Based Inflation Forecasts

The econometric evidence from the Gordon and TVI Phillips-curve specifications suggests that the output gap is not an important driver of inflation dynamics. Of course, the validity of this conclusion is conditioned on the appropriateness of the models used for describing the inflation process. In this section we provide an external check of this appropriateness for the TVI specification. Specifically, we compute inflation forecasts from the TVI specification and compare these forecasts with measures obtained from surveys of professional forecasters. To the extent that the TVI model-based forecasts are close to those obtained from surveys, it suggests that the TVI specification is a reasonable description of the evolution of inflation expectations.

We first describe how forecasts are generated from the TVI specification in equations (5)–(6). To begin, rewrite equation (5) as

$$\pi_t = z_t + \sum_{i=0}^{N} \alpha_i \begin{bmatrix} D_{t-i} \\ X_{t-i} \end{bmatrix} + \varepsilon_t, \tag{7}$$

where α_i is a vector of coefficients taken from the lag polynomials $b(L)$ and $c(L)$, and N is the lag order of these lag polynomials. Incrementing the time index in equation (7) by one quarter and taking conditional expectations yields

$$E_t[\pi_{t+1}] = E_t[z_{t+1}] + \alpha_0 E_t \begin{bmatrix} D_{t+1} \\ X_{t+1} \end{bmatrix} + \sum_{i=1}^{N} \alpha_i \begin{bmatrix} D_{t+1-i} \\ X_{t+1-i} \end{bmatrix}. \tag{8}$$

Assume that $\begin{bmatrix} D_{t+1} \\ X_{t+1} \end{bmatrix}$ can be modeled as a stationary VAR process:[7]

$$\begin{bmatrix} D_{t+1} \\ X_{t+1} \end{bmatrix} = \sum_{i=0}^{J} \beta_i \begin{bmatrix} D_{t-i} \\ X_{t-i} \end{bmatrix} + v_{t+1}. \tag{9}$$

[7] Our forecasting model for $(D_{t+1}, X_{t+1})'$ is a restricted VAR with four lags. Estimates of an unrestricted VAR, $(I - \beta(L))(D_{t+1}, X_{t+1})' = v_{t+1}$, indicated a lower triangular structure for $\beta(L)$ when the three variables are ordered as follows: (i) relative food and energy price changes, (ii) relative import price changes, and (iii) the output gap. This structure was imposed to generate our forecasts.

Then

$$E_t[\pi_{t+1}] = z_t + \alpha_0 \sum_{i=0}^{J} \beta_i \begin{bmatrix} D_{t-i} \\ X_{t-i} \end{bmatrix} + \sum_{i=1}^{N} \alpha_i \begin{bmatrix} D_{t+1-i} \\ X_{t+1-i} \end{bmatrix}. \quad (10)$$

Since $\begin{bmatrix} D_t \\ X_t \end{bmatrix}$ is assumed to be stationary, $\lim_{M\to\infty} E_t\pi_{t+M} = z_t$.[8] Thus z_t represents the long-horizon inflation forecast from the model and, in the sense of Beveridge and Nelson (1981), represents the long-run or permanent component of inflation. Likewise, $\alpha_0 \sum_{i=0}^{J} \beta_i \begin{bmatrix} D_{t-i} \\ X_{t-i} \end{bmatrix} + \sum_{i=1}^{N} \alpha_i \begin{bmatrix} D_{t+1-i} \\ X_{t+1-i} \end{bmatrix}$ is then the one-period-ahead transitory component of expected inflation.[9]

The inflation forecast error from the TVI specification is given by

$$\pi_{t+1} - E_t[\pi_{t+1}] = \omega_{t+1} + \alpha_0 \left[\begin{bmatrix} D_{t+1} \\ X_{t+1} \end{bmatrix} - \sum_{i=0}^{J} \beta_i \begin{bmatrix} D_{t-i} \\ X_{t-i} \end{bmatrix} \right] + \varepsilon_{t+1} = \omega_{t+1} + \alpha_0 v_{t+1} + \varepsilon_{t+1}. \quad (11)$$

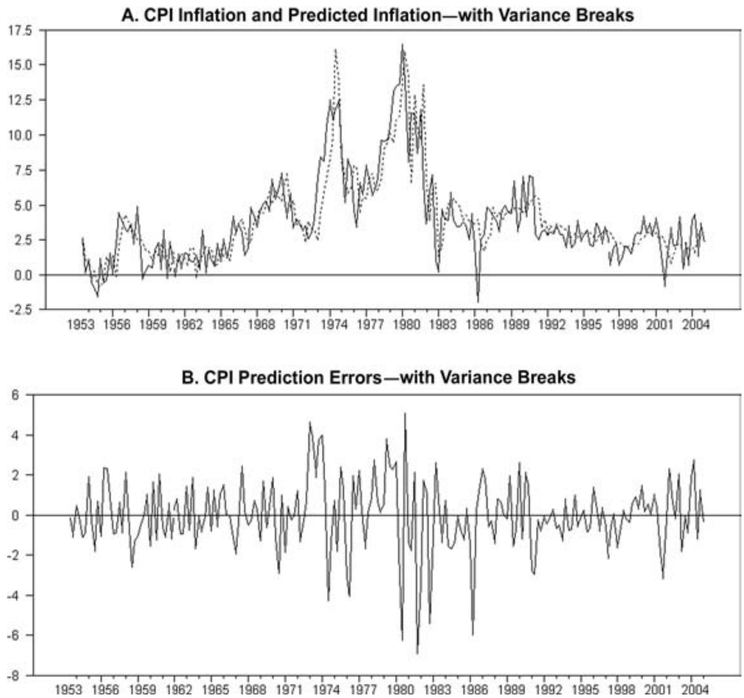
Thus unpredicted inflation is the sum of three terms: (i) the innovation to long-horizon inflation expectations, given by ω_{t+1}; (ii) the one-period-ahead forecast error for $\begin{bmatrix} D_{t+1} \\ X_{t+1} \end{bmatrix}$, given by $\alpha_0 v_{t+1}$; and (iii) the residual of the Phillips curve, given by ε_{t+1}. When $\alpha_0 = 0$ the one-period-ahead unexpected inflation is just $\pi_{t+1} - E_t[\pi_{t+1}] = \omega_{t+1} + \varepsilon_{t+1}$.

In figures 9A–11A the actual inflation rates are plotted against the one-quarter-ahead projections, $E_{t-1}[\pi_t]$, using the estimated coefficients from table 3. The lower panels of each figure (9B–11B) show the differences in the series from the

[8] This limit assumes that both D_t and X_t are mean zero, an assumption we have imposed by omitting intercepts in the VAR specification in (7). Preliminary analysis that included intercepts in (9) suggested they were statistically insignificant.

[9] By constructing multistep dynamic forecasts of $(D_{t+i}X_{t+i})'$, the entire path of the transitory component of expected inflation can be estimated.

Figure 9. CPI Inflation Predictions from TVI Model

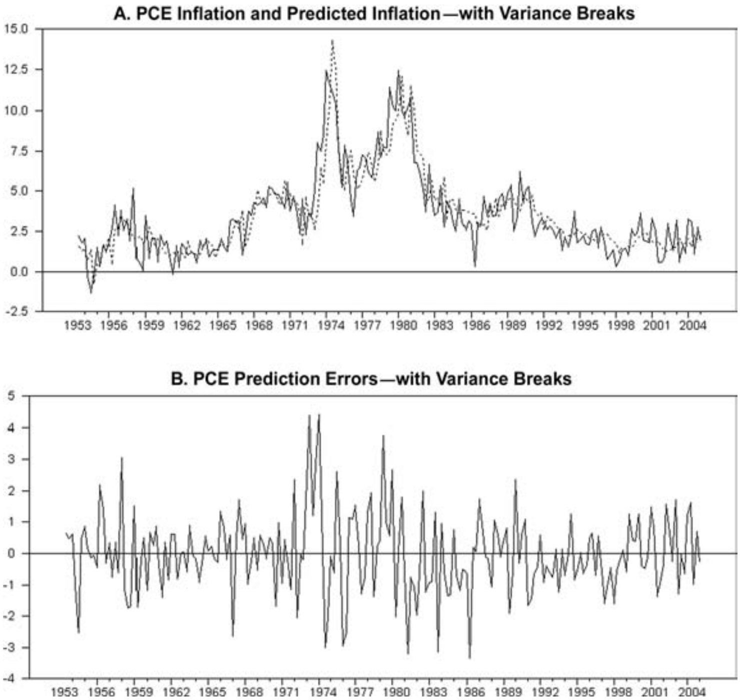


Notes: This figure contains information regarding the one-quarter-ahead inflation predictions generated by the time-varying intercept Phillips-curve specification in equations (5)–(6), where estimation was based on the sample period 1953:Q1–2005:Q1. The inflation rate is measured using the CPI. Panel A plots the actual inflation rate (solid line) against the prediction (dotted line). Panel B plots the prediction errors.

upper panels—the one-quarter-ahead inflation forecast errors.[10] The estimated autocorrelations of the computed one-quarter-ahead inflation forecast errors (not shown) are very small, indicating that there is little predictive content in the history of the forecast errors for future forecast errors.

[10]For purposes of these graphs, we incorporate the effects of the Nixon price-control dummy variables, *Nixon_On* and *Nixon_Off*. While these variables were constructed by Gordon ex post, we believe it is reasonable to assume that, at the time, individuals expected some impact on inflation in the short run of the implementation and removal of the controls.

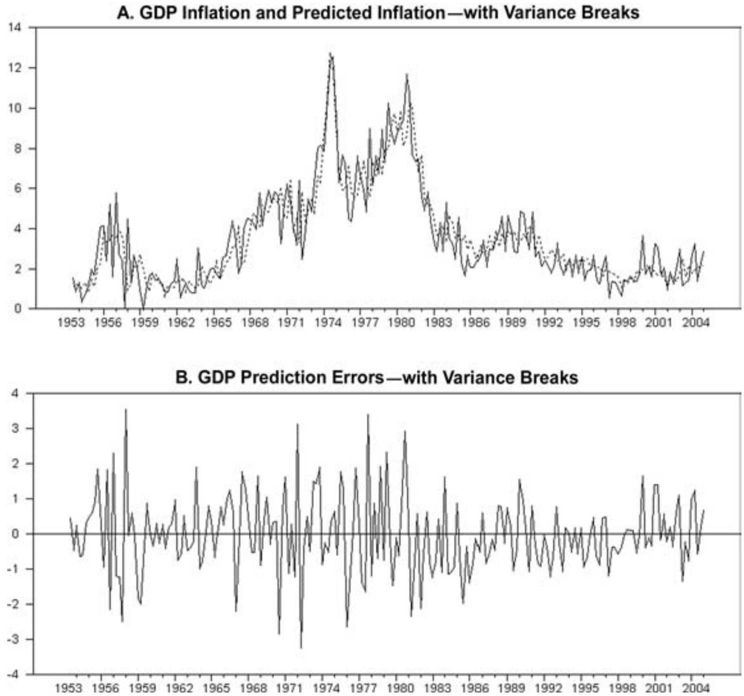
Figure 10. PCE Deflator Inflation Predictions from TVI Model



Notes: This figure contains information regarding the one-quarter-ahead inflation predictions generated by the time-varying intercept Phillips-curve specification in equations (5)–(6), where estimation was based on the sample period 1953:Q1–2005:Q1. The inflation rate is measured using the PCE index. Panel A plots the actual inflation rate (solid line) against the prediction (dotted line). Panel B plots the prediction errors.

In figure 12 we compare our estimates of the model-based one-quarter-ahead inflation forecasts with various survey measures of expected CPI and GDP deflator inflation. The inflation-forecast measure from the TVI model is indicated by the solid line in both panels of figure 12. There are two surveys that are available for CPI and GDP inflation: the one-quarter-ahead inflation forecast from the Survey of Professional Forecasters (available since 1981:Q3 for CPI inflation and 1968:Q4 for GDP inflation) and the one-quarter-ahead inflation forecast from the Blue Chip Survey (available since

Figure 11. GDP Deflator Inflation Predictions from TVI Model

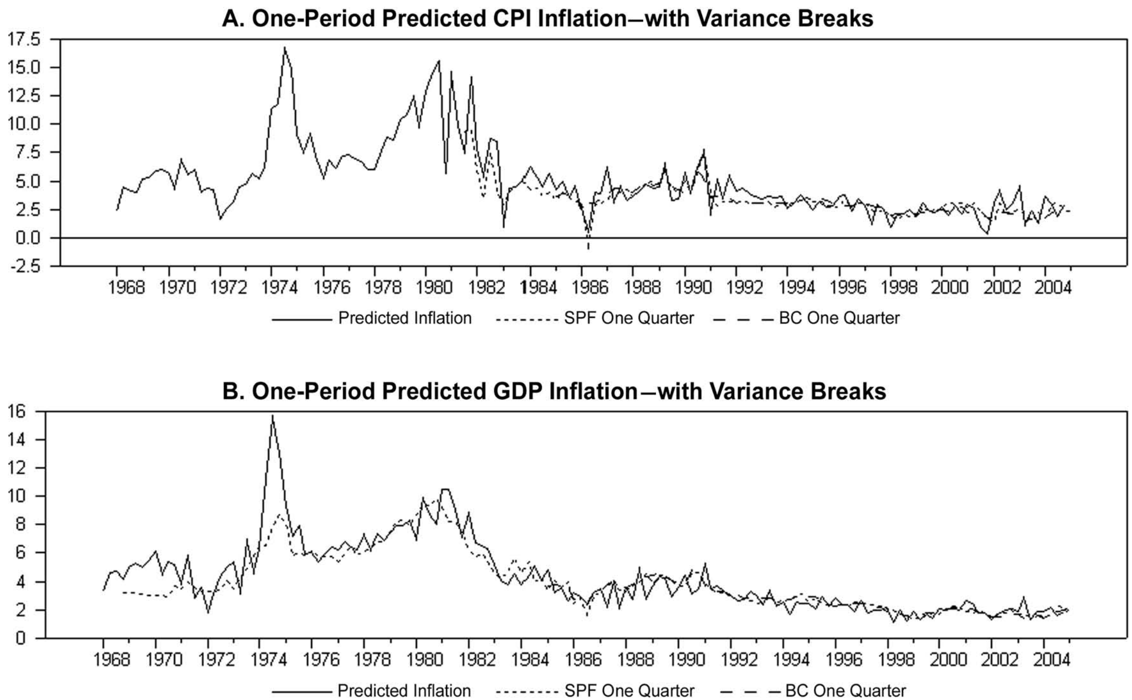


Notes: This figure contains information regarding the one-quarter-ahead inflation predictions generated by the time-varying intercept Phillips-curve specification in equations (5)–(6), where estimation was based on the sample period 1953:Q1–2005:Q1. The inflation rate is measured using the GDP index. Panel A plots the actual inflation rate (solid line) against the prediction (dotted line). Panel B plots the prediction errors.

1985:Q1 for both CPI and GDP inflation). The forecasts from the Survey of Professional Forecasters are indicated by the short-dashed line (SPF One Quarter) and the forecasts from the Blue Chip Survey are indicated by the long-dashed line (BC One Quarter) in figure 12.

For the CPI, the TVI inflation forecasts are reasonably successful at tracking the survey measures. In particular, the major spikes in the TVI inflation forecasts are mirrored in the timing, and in many

Figure 12. Inflation Predictions from TVI Model and Surveys



Notes: This figure plots the one-quarter-ahead inflation prediction generated by the time-varying intercept Phillips-curve specification in equations (5)–(6) against survey measures of inflation expectations taken from the Survey of Professional Forecasters (SPF) and Blue Chip Survey (BC). Estimation is based on the sample period 1953:Q1–2005:Q1.

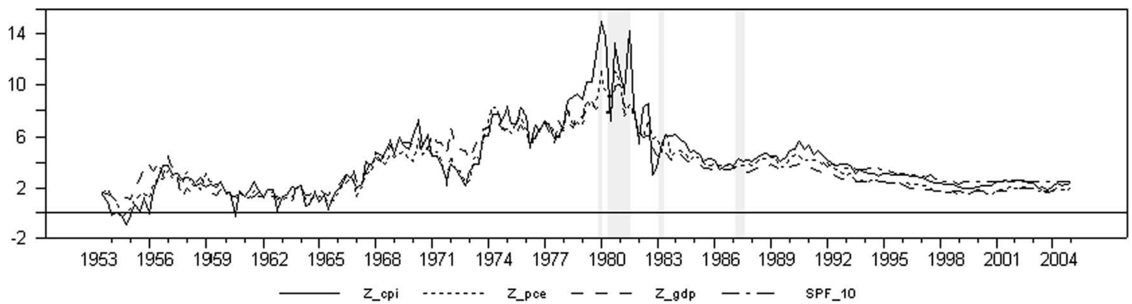
cases in the amplitude, by spikes in the SPF one-quarter measure. The BC one-quarter inflation forecasts are less volatile than the other two measures, but again the major spikes in this series mirror the timing of the major spikes in the series derived from the TVI model. The TVI inflation forecasts are less successful at tracking the survey forecasts for GDP inflation. In particular, there are substantial differences in the TVI forecast and the SPF survey forecast in the late 1960s and again in 1973. The latter period is strongly influenced by our decision to include the estimated effect of the removal of the price controls in the TVI measure of expected inflation. After 1973 the two measures track more closely, though the spikes in the TVI forecasts are not as well aligned with the survey data as is the case with the CPI inflation rate. The TVI model has the worst success at mimicking the BC survey forecast for GDP inflation. The BC survey forecast is substantially less volatile than the TVI inflation forecast, and the spikes between the two series are not particularly well aligned.[11]

The estimated time series of the time-varying intercept (the permanent component of inflation) are shown in figure 13. The series for all three inflation rates are quite similar, though the one derived from the CPI is more volatile than the other two up to the Great Moderation period. The estimates suggest that long-term expected inflation rose sharply in the late 1960s from less than 2 percent in 1964 to over 4 percent in 1968. All three series level off in the late 1960s and decline a bit in the early 1970s before the first energy shock. From 1973 until 1982 all the series trend up. From 1982 to 1985 the trend is reversed, and the series level out at around 4 percent for the remainder of the 1980s. After 1990 all the series again trend down through the mid-1990s, after which they level out at around 2 percent.

The final line (SPF_10) plotted in figure 13 is the ten-year-ahead CPI inflation forecast from the Survey of Professional Forecasters. The general trend in the long-term expected CPI inflation from the

[11]For CPI inflation, the correlation between the change in the TVI inflation forecast and the change in the survey measures is 0.56 for the SPF and 0.39 for the BC Survey. For GDP inflation, this correlation is 0.16 for the SPF and −0.11 for the BC Survey.

Figure 13. Expected Long-Term Inflation—Z and SPF



Notes: This figure plots the estimate of the time-varying intercept from the time-varying intercept Phillips-curve specification in equations (5)–(6), where estimation is based on the sample period 1953:Q1–2005:Q1. The solid, dashed, and long-dashed lines indicate estimation using CPI, PCE deflator, and GDP deflator measures of inflation, respectively. The dash-dotted line indicates a survey measure of long-horizon (ten years ahead) CPI inflation expectations taken from the Survey of Professional Forecasters (SPF).

TVI model tracks that in the survey data quite well for the period for which the latter series are available, 1991:Q4 through 2005:Q1.

5. Conclusion

We have presented evidence regarding the relative contribution of the real activity "gap" and other potential inflation drivers, such as changes in long-horizon inflation expectations and supply-shock variables, for explaining the U.S. inflation rate over the post-war period. Our results suggest that realized inflation is dominated by variation in long-horizon expected inflation, while the gap and supply-shock variables play only a very limited role. These results are robust to alternative specifications for the inflation-driving process and measures of inflation and the gap.

Our preferred model specification is one in which inflation is determined by a random-walk permanent component (which represents the long-horizon inflation expectation), a distributed lag on the real activity gap, and a distributed lag on supply-shock variables. Model-based inflation forecasts have reasonable success at tracking forecasts obtained from surveys. This suggests that our model of the inflation-driving process does a relatively good job of reproducing whatever process is driving survey measures of future inflation. Results from this model also suggest that the variance of the process that generates changes in long-horizon expected inflation has changed over time. Interestingly, this variance has become very small over the last ten years of the sample, suggesting that long-horizon expected inflation has become much better "anchored" in the past decade.

Taken together, the evidence presented here suggests that the key to understanding the inflation process is to understand what drives changes in long-horizon inflation expectations. To this end, further research focused on attempting to relate these changes to "news" could prove especially fruitful.

References

Atkeson, A., and L. E. Ohanian. 2001. "Are Phillips Curves Useful for Forecasting Inflation?" *Quarterly Review* (Federal Reserve Bank of Minneapolis) 25 (1): 2–11.

Beveridge, S., and C. R. Nelson. 1981. "A New Approach to Decomposition of Economic Time Series into Permanent and Transitory Components with Particular Attention to Measurement of the Business Cycle." *Journal of Monetary Economics* 7 (2): 151–74.

Clark, T. E., and M. W. McCracken. 2006. "The Predictive Content of the Output Gap for Inflation: Resolving In-Sample and Out-of-Sample Evidence." *Journal of Money, Credit, and Banking* 38 (5): 1127–48.

Gordon, R. J. 1982. "Inflation, Flexible Exchange Rates and the Natural Rate of Unemployment." In *Workers, Jobs and Inflation*, ed. M. N. Baily, 88–152. Washington, DC: Brookings Institution.

———. 1997. "The Time-Varying NAIRU and Its Implications for Economic Policy." *Journal of Economic Perspectives* 11 (1): 11–32.

———. 1998. "Foundations of the Goldilocks Economy: Supply Shocks and the Time-Varying NAIRU." *Brookings Papers on Economic Activity* 29 (2): 297–346.

Gürkaynak, R. S., B. Sack, and E. Swanson. 2005. "The Sensitivity of Long-Term Interest Rates to Economic News: Evidence and Implications for Macroeconomic Models." *American Economic Review* 95 (1): 425–36.

Kim, C.-J., and C. R. Nelson. 1999. "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle." *The Review of Economics and Statistics* 81 (4): 608–16.

Kozicki, S., and B. Hoffman. 2004. "Rounding Error: A Distorting Influence on Index Data." *Journal of Money, Credit, and Banking* 36 (3): 319–38.

Levin, A., and J. Piger. 2002. "Is Inflation Persistence Intrinsic in Industrial Economies?" Federal Reserve Bank of St. Louis Working Paper No. 2002–023.

———. 2005. "Bayesian Model Selection for Structural Break Models." Mimeo.

McConnell, M. M., and G. Pérez-Quirós. 2000. "Output Fluctuations in the United States: What Has Changed Since the Early 1980s?" *American Economic Review* 90 (5): 1464–76.

Orphanides, A., and S. van Norden. 2005. "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time." *Journal of Money, Credit, and Banking* 37 (3): 583–601.

Rudebusch, G. D., and T. Wu. 2008. "A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy." *Economic Journal* 118 (530): 906–26.

Stock, J. H., and M. W. Watson. 1999. "Forecasting Inflation." *Journal of Monetary Economics* 44 (2): 293–335.

———. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit, and Banking* 39 (s1): 3–33.

Leadership in Groups: A Monetary Policy Experiment*

Alan S. Blinder[a] and John Morgan[b]
[a]Princeton University
[b]University of California, Berkeley

This paper studies monetary policy decision making by committee, using an experimental methodology. In an earlier paper (Blinder and Morgan 2005), we found that groups not only outperformed individuals, but they also took no longer to reach decisions. We successfully replicate those results here. Next, we find little difference between the performances of four-person and eight-person groups; the larger groups outperform the smaller groups by a very small (and often insignificant) margin. Third, and most surprisingly, we find no evidence of superior performance by groups that have designated leaders. Possible reasons for that strongly counterintuitive finding are discussed.

JEL Codes: C92, E58.

1. Introduction and Motivation

The transformation of monetary policy decisions in most countries from individual decisions to group decisions is one of the most notable developments in the recent evolution of central banking (Blinder 2004, ch. 2). In an earlier paper (Blinder and Morgan 2005), we ran an experiment in which Princeton University students, acting as ersatz central bankers, made monetary policy decisions both as individuals and in groups. Those experiments yielded two main findings:

*We are grateful to Jennifer Brown, Jae Seo, and Patrick Xiu for fine research assistance and to the National Science Foundation and Princeton's Center for Economic Policy Studies for financial support. We also acknowledge extremely helpful comments from Petra Geraats, Petra Gerlach-Kristen, Jens Grosser, Helmut Wagner, a referee, and seminar participants at Princeton, the International Monetary Fund, and the National Bureau of Economic Research.

(i) Groups made better decisions than individuals, in a sense to be made precise below.
(ii) Groups took no longer to reach decisions than individuals did.[1]

The first finding was not a big surprise, given the previous literature on group versus individual decision making (most of it from disciplines other than economics). But we were frankly stunned by the second finding. Like seemingly everyone, we believed that groups moved more slowly than individuals. A subsequent replication with students at the London School of Economics (Lombardelli, Proudman, and Talbot 2005) verified the first finding but did not report on the second one.

This paper replicates our 2005 findings using the same experimental apparatus, but with students at the University of California, Berkeley. That the replication is successful bolsters our confidence in the Princeton results. But that is not the focus of this paper. Instead, we study two important issues that were deliberately omitted from our previous experimental design.

The first pertains to *group size*. In the Princeton experiment, every monetary policy committee (MPC) had five members—precisely (and coincidentally) the size that Sibert (2006) subsequently judged to be optimal. Lombardelli, Proudman, and Talbot (2005), following our lead, also used committees of five. But real-world monetary policy committees vary greatly in size, so it seems important to compare the performance of small versus large groups. Revealed-preference arguments offer little guidance in this matter, since real-world MPCs range in size from three to twenty-one, with the European Central Bank (ECB) headed even higher. In this paper, we study the size issue by comparing the experimental performances of groups of four and eight.[2]

[1]In both our 2005 paper and the present one, "time" is measured by the amount of *data* required before the individual or group decides to change the interest rate—not by the number of ticks of the clock. Our reason was (and remains) simple: this is the element of time lag that is relevant to monetary policy decisions; no one cares about how many hours the committee meetings last.

[2]The reason for choosing even-numbered groups will be made clear shortly. Our "large" groups ($n = 8$) are still small compared with, e.g., the ECB or the Federal Reserve. This size was more or less dictated by the need to recruit large

The second issue pertains to *leadership* and is the unique aspect of the research reported here. Both our Princeton experiment and Lombardelli, Proudman, and Talbot's replication treated all members of the committee equally. But every real-world monetary policy committee has a designated leader who clearly outranks the others. At the Federal Reserve, that leader is the "chairman"; at the ECB, he is the "president"; and at the Bank of England and many other central banks, he or she is the "governor." Indeed, we are hard-pressed to think of *any* committee, in *any* context, that does *not* have a well-defined leader. Juries come close, but even they have foremen. Observed reality, therefore, strongly suggests that groups need leaders in order to perform well. But is it true? That is the main question that motivates this research.

Consider leadership on MPCs in particular. While all MPCs have designated leaders, the leader's authority varies greatly. The Federal Open Market Committee (FOMC) under Alan Greenspan (but not under Ben Bernanke) was at one extreme; it was what Blinder (2004, ch. 2) called an *autocratically collegial* committee, meaning that the chairman came close to dictating the committee's decision. This tradition of strong leadership did not originate with Greenspan. Paul Volcker's dominance was legendary, and Chappell, McGregor, and Vermilyea (2005, ch. 7) estimated econometrically that Arthur Burns's views on monetary policy carried about as much weight as those of all other FOMC members combined. At the other extreme, the Bank of England's MPC is what Blinder (2004) called an *individualistic* committee—one that reaches decisions (more or less) by true majority vote. Governor Mervyn King has even allowed himself to be outvoted, partly in order to make this point. In between these poles, we find a wide variety of *genuinely collegial* committees, like the ECB Governing Council, that strive for consensus. Some of these committees are led firmly; others are led only gently.

The scholarly literature on group decision making, which comes mostly from psychology and organizational behavior, offers relatively little guidance on what to expect. And only a small portion of it is experimental. As a broad generalization, our quick review of the literature led us to expect to find some positive effects of leadership on

numbers of subjects. With groups of four and eight, we needed 252 subjects in all.

group performance—which is the same prior we had before reviewing the literature. But it also led to some doubts about whether intellectual ability is a key ingredient in effective leadership (Fiedler and Gibson 2001). Instead, the literature suggests that gains from group interaction may depend more on how well the leader encourages other members of the group to contribute their opinions frankly and openly (Maier 1970, Blades 1973, and Edmondson 1999). In an interesting public goods experiment, Guth et al. (2004) found that stronger leadership produced better results, although the leaders in that experiment were selected randomly. We did not find any relevant evidence on whether leadership effects are greater in larger or smaller groups.

With these two issues—group size and leadership—in mind, we designed our experiment with four treatments, running ten or eleven sessions with each treatment:

(i) four-person groups with no leader, hereafter denoted $\{n = 4$, no leader$\}$
(ii) four-person groups with a leader $\{n = 4$, leader$\}$
(iii) eight-person groups with no leader $\{n = 8$, no leader$\}$
(iv) eight-person groups with a leader $\{n = 8$, leader$\}$

We summarize our results briefly here because they will be understood far better after the experimental details are explained. First, we successfully replicate our Princeton results, at least qualitatively: groups perform better than individuals, and they do not require more "time" to do so. Second, we find little difference between the performance of four-person and eight-person groups; the larger groups outperform the smaller groups by a very small (and often insignificant) margin. Third, and most important, we find no evidence of superior performance by groups that have designated leaders. Groups without such leaders do as well as or better than groups with well-defined leaders. This is a surprising finding, and we will speculate on some possible reasons later.

The rest of the paper is organized as follows. Section 2 describes the experimental setup, which in most respects is exactly the same as in Blinder and Morgan (2005). Section 3 briefly presents results comparing group and individual performance that mostly replicate those of our Princeton experiment. Sections 4 and 5 focus on the data generated by decision making in groups, presenting new results on the effects of group size and leadership, respectively. Then section 6 summarizes the conclusions.

2. The Experimental Setup[3]

Our experimental subjects were Berkeley undergraduates who had taken at least one course in macroeconomics. We brought them into the Berkeley Experimental Social Sciences Lab (Xlab) in groups of either four or eight, telling them only that they would be playing a monetary policy game. Except by coincidence, the students did not know one another beforehand. Each computer was programmed with the following simple two-equation macroeconomic model—exactly the same one used in the Princeton experiment—with parameters chosen to resemble the U.S. economy:

$$\pi_t = 0.4\pi_{t-1} + 0.3\pi_{t-2} + 0.2\pi_{t-3} + 0.1\pi_{t-4} - 0.5(U_{t-1} - 5) + w_t \tag{1}$$

$$U_t - 5 = 0.6(U_{t-1} - 5) + 0.3(i_{t-1} - \pi_{t-1} - 5) - G_t + e_t. \tag{2}$$

Equation (1) is a standard accelerationist Phillips curve. Inflation, π, depends on the deviation of the lagged unemployment rate from its presumed natural rate of 5 percent, and on its own four lagged values, with weights summing to one. The coefficient on the unemployment rate is chosen roughly to match empirically estimated Phillips curves for the United States.

Equation (2) can be thought of as an IS curve with the unemployment rate, U, replacing real output (via Okun's Law). Unemployment tends to rise above (or fall below) its natural rate when the *real* interest rate, $i - \pi$, is above (or below) its "neutral" value, which is also set at 5 percent. (Here i is the nominal interest rate.) But there is a lag in the relationship, so unemployment responds to the real interest rate only gradually. Like real-world central bankers, our experimental subjects control only the *nominal* interest rate, not the *real* interest rate.

[3]This section overlaps substantially with section 1.1 of Blinder and Morgan (2005) but omits some of the detail presented there.

The G_t term in (2) is the shock to which our student monetary policymakers are supposed to react. It starts at zero and randomly changes *permanently* to either +0.3 or −0.3 sometime during the first ten periods of play. Readers can think of G as representing government spending or any other shock to aggregate demand. As is clear from (2), a change in G changes U by precisely the same amount, but in the opposite direction, on impact. Then there are lagged responses, and the model economy eventually converges back to its natural rate of unemployment. Because of the vertical long-run Phillips curve, any constant inflation rate paired with $U = 5\%$ can be an equilibrium.

We begin each round of play with inflation at 2 percent—which is also the central bank's target rate (see below). Thus, prior to the shock (i.e., when $G = 0$), the model's steady-state equilibrium is $U = 5$, $i = 7$, $\pi = 2$. As is apparent from the coefficients in equation (2), the shock changes the neutral real interest rate from 5 percent to either 6 percent or 4 percent *permanently.* Our subjects—who do *not* know this—are supposed to detect and react to this change, presumably with a lag, by raising or lowering the nominal interest rate accordingly.

Finally, the two stochastic shocks, e_t and w_t, are drawn independently from uniform distributions on the interval $[-.25, +.25]$.[4] Their standard deviations are roughly half the size of the G shock. This sizing decision, we found, makes the fiscal shock relatively easy to detect—but not too easy.

Lest our subjects had forgotten their basic macroeconomics, the instructions remind them that raising the interest rate lowers inflation and raises unemployment, while lowering it does the reverse, albeit with a lag.[5] In the model, monetary policy affects unemployment with a one-period lag and inflation with a two-period lag; but students are not told that. Nor are they told anything else about the model's specification. They *are* told that the demand shock will occur at a random time that is equally likely to be any of periods 1 through 10. But they are told neither the magnitude of this shock, nor its direction, nor whether it is permanent or temporary.

[4] The distributions are uniform, rather than normal, for programming convenience.

[5] The instructions are provided in the appendix.

Doubtless, this little model economy is far simpler than the actual economies that real-world central bankers try to manage. However, *to the student subjects*, who do not know anything about the model, we believe this setup poses perplexities that are comparable to those facing real-world central bankers, who are trying to stabilize a much more complex system (e.g., one that includes expectational effects) but who also know much more, have far more experience, and have abundant staff support. For example, our experimental subjects do not know the transmission mechanism, the lag structure, whether the price equation is forward looking or backward looking, and so on. Nor do they have the benefit of staff forecasts or analyses.

Despite the model's seeming simplicity, stabilizing it can be tricky in practice. Because of the unit root apparent in equation (1), the model diverges from equilibrium when perturbed by a shock—unless it is stabilized by monetary policy. But lags and modest early-period effects combine to make the divergence from equilibrium pretty gradual and hence less than obvious at first. Once unemployment and inflation start to "run away from you," it can be difficult to get them back on track. Furthermore, it is not easy to distinguish between the permanent G shock and the transitory e and w shocks that add "noise" to the system. Indeed, the subjects do not even know that the G shock is permanent while the others are i.i.d.

Each play of the game proceeds as follows. We start the system in steady-state equilibrium at the values mentioned above. The computer then selects values for the two random shocks and displays the first-period values of U and π, which are typically quite close to the target values ($U = 5\%$, $\pi = 2\%$), on the screen for the subjects to see. In each subsequent period, new random values of e_t and w_t are drawn, thereby creating statistical noise, and the lagged variables that appear in equations (1) and (2) are updated. At some random time, unknown to subjects, the G shock occurs. The computer calculates U_t and π_t each period and displays them on the screen, where all past values are also shown. Subjects are then asked to choose an interest rate for the next period, and the game continues for twenty such periods. Students are told to think of each period as a quarter, so the simulation covers "five years." Each five-year run of the game is different because of different random draws.

No time pressure is applied; subjects are permitted to take as much clock time as they wish to make each decision. As noted above, the concept of time that interests us is the *decision lag*: the amount of *new data* the decision maker insists upon before changing the interest rate. In the real world, data flow in unevenly over calendar time; in our experiment, subjects see exactly one new observation on unemployment and inflation each period. So when we say later that one type of decision-making process "takes longer" than another, we mean that more *data* (not more *minutes*) are required.

To rate the quality of their performances, and to reward subjects accordingly, we tell students that their score for each quarter is

$$s_t = 100 - 10|U_t - 5| - 10|\pi_t - 2|, \tag{3}$$

and the score for the entire game (henceforth, S) is the (unweighted) average of s_t over the twenty quarters. We use an absolute-value function instead of the quadratic loss function that is ubiquitous in research on monetary policy (and elsewhere) because quadratics are too hard for subjects—even Princeton and Berkeley students—to calculate in their heads. Notice also that the coefficients in equation (3) scale the scores into percentages, which gives them a natural, intuitive interpretation. Thus, e.g., missing the unemployment target by 0.8 (in either direction) and the inflation target by 1.0 results in a score of $100 - 8 - 10 = 82$ (percent) for that period.[6] At the end of the session, scores are converted into money at the rate of 25¢ per percentage point. Subjects typically scored 80–84 percent of the possible points, thus earning about \$20–\$21.

One final detail needs to be mentioned. To deter excessive manipulation of the interest rate (which we observed in testing the apparatus in dry runs), we charge subjects a fixed cost of 10 points each time they change the rate of interest, regardless of the size of the change.[7] Ten points is a small charge; averaged over a twenty-period game, it amounts to just 0.5 percent of the total potential score. But we found it to be large enough to deter most of the excessive fiddling with interest rates. Analogously, researchers who try to derive

[6]The unemployment and inflation data are always rounded to the nearest tenth. So students see, e.g., 5.8 percent, not, say, 5.83 percent.

[7]To keep things simple, only integer interest rates are allowed.

the Federal Reserve's reaction function from the minimization of a quadratic loss function find that they must add something like $k(i_t - i_{t-1})^2$ to the loss function in order to fit the data. Without that term, interest rates turn out to be far more volatile than they are in practice.[8]

The sessions are played as follows. Either four or eight students enter the lab and are read detailed instructions, which they are also given in writing. (See the appendix.) In the case of sessions with a designated leader, the instructions tell them, among other things, that the person earning the highest score while playing alone in part 1 of the experiment will be designated the "leader" (the term we use) of the group for part 2—where he or she will be rewarded with a doubled score.

Subjects are then allowed to practice with the computer apparatus for five minutes, during which time they can ask any questions they wish. At the end of the practice period, each machine is reinitialized, and each student is instructed to play twelve rounds of the game (each lasting twenty "quarters") *alone*—without communicating in any way with the other subjects. Once all the subjects have completed twelve rounds of individual play, the experimenter calls a halt to part 1 of the experiment.

In part 2, the same students gather around a single large screen to play the same game twelve times *as a group*. It is here that the sessions with and without leaders differ. In leaderless sessions, the rules are exactly the same as in individual play, except that students are now permitted to communicate freely with one another—as much as and in any way they please. Everyone in the group is treated alike, and each subject receives the group's common score.

In sessions with a designated leader, the experimenter begins by revealing who earned the highest score in part 1, and that student becomes the leader for part 2.[9] Thus, the criterion for electing leaders is purely intellective: the skill of an individual at ersatz monetary policymaking. Since the group will perform the identical task, this selection principle would seem a natural one.

[8] See, e.g., Rudebusch (2001).

[9] On average, that student scored 10.8 points higher than the others in the group during part 1 of the experiment, a sizable difference. But students were not told the leader's score.

Table 1. The Flow of the Experiment

Instructions
Practice Rounds (no scores recorded)
Part 1: Twelve rounds played as individuals
Part 2: Twelve rounds played as a group (with or without a leader)
Part 3: Twelve rounds played as individuals
Students are paid by check and leave

The meaning of leadership in the experiment is threefold: First, the leader is responsible for communicating (verbally) the group's decision to the experimenter—which makes the leader pivotal to the discussion. Second, the leader faces stronger incentives: his or her score in part 2 is *double* that of the other subjects. Third, the leader gets to break any tie vote—which is why we use even-numbered groups.[10] While we recognize that the experimental setup still allows limited scope for leadership, we judged that this was about all we could do in a laboratory setting with $1^1/_2$ hours of experimental time. We return to this issue later.

After twelve rounds of group play, the subjects return to their individual computers for part 3, in which they play the game another twelve times alone, with no communication with the others. For future reference, table 1 summarizes the flow of each session.

A typical session (of 36 rounds of the game) lasted about 90 minutes, and we ran 42 sessions in all, amounting to 252 total subjects. (No subject was permitted to play more than once.) Each of the 21 four-person sessions *should have* generated 24 individual rounds of play per subject, or $21 \times 4 \times 24 = 2,016$ in all, plus 12 group rounds per session, or 252 in all. Each of the 21 eight-person sessions *should have* generated twice as many individual observations (hence 4,032 in total), plus another 252 group observations. Thus we have a plethora of data on individual performance but a relative paucity of data on group performance. Since a small number of observations were lost due to computer glitches, table 2 displays the exact number of observations we actually generated for

[10]In principle, the tie-breaking privilege should be worth more in groups of four than in groups of eight. In practice, however, ties were rare.

Table 2. Number of Observations for Each Treatment

	Number of Sessions	Individuals	Groups
$n = 4$, no leader	10	960	120
$n = 4$, leader	11	1,032	132
$n = 8$, no leader	10	1,885	120
$n = 8$, leader	11	2,112	132
All Treatments	42	5,989	504

each treatment. Most of this paper concentrates on our new findings on the behavior of ersatz monetary policy committees—the 504 experimental observations listed in the far right column of table 2.

3. Groups versus Individuals: A Replication

We turn first, and very briefly, to the 5,989 observations on individual performance and, especially, to the comparisons between groups and individuals that were the focus of Blinder and Morgan (2005). The results here are easy to summarize: for the most part, our new results with the Berkeley sample replicate what we found earlier with the Princeton sample.[11]

To begin with, we found in our Princeton experiment that groups (which were all of size five) turned in better average performances than did individuals. Specifically, the average group score was 88.3, while the average individual score was 85.3. The difference of 3 points, or 3.5 percent, was highly significant. If we merge all four of our group treatments in the Berkeley experiment, the average group score is 86.6 versus an average individual score of 81.1. Again, groups do better, but here their advantage is 5.5 points, or 6.8 percent—almost twice as large as in the Princeton experiment. This performance gap is also highly significant ($t = 11.2$).

The following regression confirms that this quantitative (but not qualitative) difference between the two experimental results is significant. Clustering by session to produce robust standard errors yields

[11] However, the Princeton and Berkeley samples have different statistical properties, including both first and second moments, which is why we abandoned our original idea of simply merging the two samples.

the following linear regression, with standard errors in parentheses and absolute values of t-ratios under that:[12]

$$\begin{array}{lllll} S_i = 85.27 & + \; 3.02\, GP_i & -\, 4.18\; BERK_i & + \, 2.50\big(GP_i \, {}^{*}BERK_i\big) & \\ \quad (0.37) & (0.57) & (0.55) & (0.75) & \\ \quad t = 231.8 & t = 5.4 & t = 7.6 & t = 3.4 & \\ & & R^2 = 0.027 & N = 8,893 & \quad (4) \end{array}$$

where GP and $BERK$ are dummy variables associated with observations that occurred when the game was played as a *group* and by *Berkeley* students, respectively. The coefficient estimates, all of which are significant at the 1 percent level, reveal that Berkeley students perform worse than Princeton students but improve more from group interaction. We do not have a ready explanation for this difference, but we do note that Lombardelli, Proudman, and Talbot (2005, 194) found that weaker players improved more over the course of their entire experiment—spanning both group and individual play.

This finding, plus some others to be mentioned below, suggests a systematic pattern: weaker players may gain more from exposure to group play. To investigate this phenomenon a bit further, we disaggregated both our Berkeley and Princeton samples to see whether the *increase* in scores from part 1 (individual play) to part 2 (group play) correlated *negatively* with the part 1 scores. That is, do weaker players gain more from working in groups? To assess ability, we employ a natural, high-quality control for the skill of each group—namely, the average score of the group's members *prior to* group play, i.e., in part 1. We call this variable A or *Ability*. Regressing the mean score of a group over its twelve repetitions (*Gmean*) on A leads to

$$\begin{array}{llll} Gmean_i = & 56.77 \; + & 0.386\, A_i \quad R^2 = 0.320 \quad N = 351 & \\ & (8.90) & (0.11) & \\ & t = 6.38 & t = 3.50 & \quad (5) \end{array}$$

[12] Clustering by session allows for the possibility of autocorrelation and heteroskedasticity for observations generated in a given session (i.e., by the same group of individuals). See White (1980).

Notice that the coefficient on the average individual score is considerably below unity, implying that $Gmean - A$, the improvement in group play, is decreasing in A. Thus, consistent with the findings of Lombardelli, Proudman, and Talbot (2005), we find that weaker players improve more than do stronger players from group interaction.

The next question pertains to the decision-making lag. How much time elapses, on average, between the shock and the monetary policy reaction to it? And do groups display systematically longer lags than individuals? Remember, the most surprising result from our original Princeton experiment was that groups were *not* slower; in fact, they were slightly faster, though not significantly so. Approximately the same is true in our Berkeley experiment. The mean lags before the *first* interest rate change are essentially identical (roughly 3.3 "quarters") in both group and individual play.

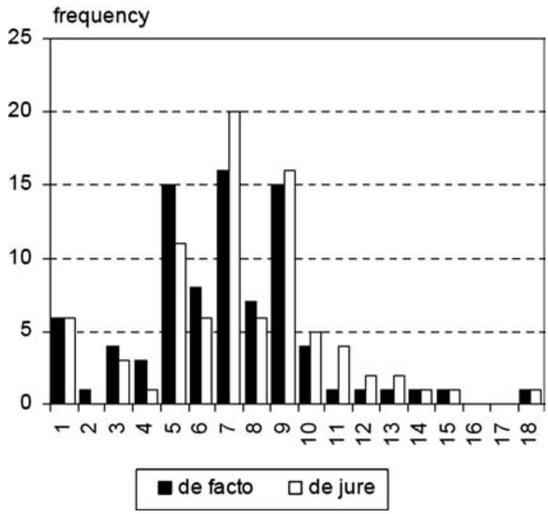
To investigate this question, we create the dependent variable *Lag*, defined as the number of quarters that elapse between the shock (the increase or decrease in G) and the committee's *first* interest rate change. Regression (6) estimates the same specification as (4), but with *Lag* replacing S as the dependent variable:

$$\begin{array}{llllll} Lag_i = & 2.45 & - \; 0.15\, GP_i & + \; 0.75\, BERK_i & + \; 0.12\, GP_i * BERK_i & \\ & (0.23) & (0.21) & (0.28) & (0.30) & \\ & t = 10.7 & t = 0.7 & t = 2.7 & t = 0.4 & \\ & & & & R^2 = 0.007 \quad N = 8{,}893 & \end{array} \tag{6}$$

This regression shows that groups take about the same amount of time as individuals to reach a decision, as we found before. (The F-test for omitting the two GP variables has a p-value of 0.69) It also shows that Berkeley students playing as individuals move more slowly (by approximately 0.75 "quarters") than Princeton students.

This demonstrated ability to replicate our earlier results enhances confidence in the experimental apparatus. So we turn now to the two new questions, which pertain to group size and leadership. Since the two issues are largely orthogonal, we treat them separately at first. Later (cf. table 4), we will show that interaction effects between group size and leadership are negligible and statistically insignificant.

Figure 1. Distribution of MPC Size in the Sample



Source: Erhart and Vasquez-Paz (2007)

4. Are Larger Groups More Effective Than Smaller Groups?

The title of our 2005 paper asked metaphorically, are two heads better than one? We now ask—literally—whether eight heads are better than four; i.e., do smaller ($n = 4$) or larger ($n = 8$) groups perform better in conducting simulated monetary policy?

As an empirical matter, most real-world MPCs cluster in the five- to ten-member range, with some smaller and some larger.[13] The most recent study of committee size, by Erhart and Vasquez-Paz (2007), finds that both the mean and median sizes of committees are around seven members. Figure 1, which is taken from that paper, also illustrates that the distribution of committee size is asymmetric—with a small but long right-hand tail.[14] In addition, it

[13]See Mahadeva and Sterne (2000).

[14]Erhart and Vasquez-Paz (2007) distinguish between de facto and de jure size, which do not always match up. Figure 1 shows both.

can be seen that committees with odd numbers of members are far more common than committees with even numbers. So our larger committees are somewhat typical of real-world MPCs, while our smaller committees are clearly on the small side. But does group size matter at all?

To focus on size effects, we begin by pooling the data from sessions with and without designated leaders—a pooling that our subsequent results say is legitimate. Initially, we do not control for the skill levels of the members of the group either. Simply regressing the average game score (the variable S defined above) for each of the 504 group observations on a dummy for the size of the group, and clustering by session to produce robust standard errors, yields the simple linear regression shown in column 1 of table 3, with standard errors in parentheses and asterisks indicating significance levels. The "large-group dummy" is defined to be 1 for eight-person groups and 0 for four-person groups. Thus, the regression suggests a small positive effect of larger group size—a score 2.3 points higher for the larger groups—which is significant if you are not too fussy about significance levels (the p-value is 0.067).

However, larger groups might simply have drawn, on average, more highly skilled individuals than did smaller groups. So it seems advisable to control for the abilities of the various members of the group. Fortunately, we have a natural, high-quality control for ability: the average score of all the members of the group *prior to* their exposure to group play, i.e., in part 1 of the experiment. This is the variable A introduced above, and we use both it and its square as controls for skill in the column 2 regression. Notice the huge jump in R^2—the $Ability$ variable has high explanatory power.[15]

Column 2 reveals that controlling for differences in the average ability of members of the larger groups reduces the estimated difference in the performance of large versus small groups by over 40 percent—to just 1.3 points. However, even after accounting for the ability of group members, larger groups perform significantly better (p-value = 0.08) than smaller groups.

[15]When the same regression is estimated by ordinary least squares, the coefficients are almost identical, but the standard errors are roughly half of those in column 2—indicating that clustering matters.

Table 3. Regression Results on Group Size

Dependent Variable→	(1) Score	(2) Score	(3) Score	(4) Score	(5) *Lag*	(6) *Correct*	(7) *Frequency*
Large-Group Dummy	2.280*	1.292*	1.033	1.031	−0.021	−0.011	−0.269*
	(1.211)	(0.723)	(0.655)	(0.657)	(0.429)	(0.038)	(0.150)
Ability		9.628***	7.027***	7.077***	−2.331**	0.006	−0.133
		(3.276)	(2.421)	(2.629)	(0.912)	(0.114)	(0.366)
$Ability^2$		−0.060***	−0.044**	−0.044**	0.014**	0.000	0.001
		(0.021)	(0.016)	(0.017)	(0.006)	(0.001)	(0.002)
Best			2.023	1.981			
			(1.857)	(1.902)			
$Best^2$			−0.010	−0.010			
			(0.012)	(0.012)			
Group Standard Deviation				0.020			
				(0.156)			
Constant	85.479***	−300.528**	−293.160***	−293.370***	97.332***	0.437**	6.068
	(1.063)	(124.108)	(85.630)	(86.630)	(33.718)	(4.256)	(13.616)
No. of Observations	504	504	504	504	504	504	504
R^2	0.03	0.24	0.26	0.26	0.07	0.01	0.03

Notes: Robust standard errors clustered by session are in parentheses. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively. "Large-Group Dummy" equals 1 if the group contained eight members. "*Ability*" is the average score of all members of a group during part 1 of the experiments. "*Best*" is the highest average score attained by a member of a group during part 1. "Group Standard Deviation" is the standard deviation of average part 1 scores.

The estimated quadratic in *Ability*, by the way, carries an interesting and surprising implication: that the contribution of individual ability to group performance peaks at $A = 80.7$ points, which is only a few points above the average part 1 score of 77.4 points. After that, too many good cooks seem to spoil the broth.

The negative slope beyond $A = 80.7$ is, however, an artifact of the inflexible quadratic functional form. When we estimate instead a freer functional form (such as a spline) that allows the relationship between S and A to flatten out beyond, say, $A = 80$, we get essentially a zero (rather than a negative) slope for high values of A. Still, it is surprising that groups reap no further rewards from the abilities of their members once A exceeds a fairly modest level (approximately 80). But this is a robust finding that survives experiments with several functional forms.[16]

Let us now return to why larger groups perform (slightly) better than smaller groups. One possibility is that a group's decisions are dominated by its most skilled player.[17] Larger groups will, on average, have better "best players" than smaller groups simply because the first-order statistic for skill will, on average, be higher when $n = 8$ than when $n = 4$. To see whether that factor might be empirically important in these data, we added both the average score of the group's best player (*Best*) and its square in the regression to get the regression reported in column 3. We see that the effect of larger group size is reduced by about 20 percent, and it is now no longer significant at even the 10 percent level ($p = 0.12$).

The explanatory power of the *Best* variables is modest, however. Neither *Best* nor $Best^2$ is statistically significant on its own, and the estimated coefficients are small compared with those of the A variables. Moreover, adding *Best* and $Best^2$ raises R^2 by only 0.026.[18] However, an F-test of the joint hypothesis that the coefficients on

[16] The surprising thing is not that there are diminishing returns to group size, which can be rationalized in many ways, but that marginal returns seem to get to zero so quickly.

[17] Several colleagues *assured* us that this would be the case in our first experiment. But we tested and rejected the hypothesis in Blinder and Morgan (2005).

[18] Surprisingly, the individual score of the *second-best* player turns out to have more explanatory power for the group's performance. We have no ready explanation for this finding and treat it as a fluke. Regardless, the results on group size are not qualitatively affected under this alternative specification.

both variables are 0 strongly rejects that hypothesis ($F = 30.9$, $p = 0.00$).[19] Thus, the evidence suggests that the fuller specification (column 3) is preferred, but that the influence of the best player is modest—a point to which we shall return in considering the effects of leadership.

Next, we ask whether heterogeneity of the members of the group, as measured by skill differences across players, improves group performance. We measure heterogeneity by introducing a new variable in column 4: the standard deviation of the average part 1 scores obtained by the (four or eight) members of the group.[20] Comparing columns 3 and 4 shows that adding this variable has essentially no effect on the regression. Thus heterogeneity does not seem to matter.

4.1 How Do Larger Groups Outperform Smaller Groups?

Having shown that larger groups (barely) outperform smaller groups, the next question is, how do they do it? To determine whether a shorter or longer decision-making lag is the source of the advantage for large groups, we regress the variable *Lag* defined above on the dummy for groups of size eight and the ability controls mentioned above, clustering by session as usual. The result is the regression in column 5 of table 3, which indicates no difference between the two group sizes in terms of speed of decision making. (The p-value of the coefficient of the dummy is 0.58.) Differences in ability are again significant, with groups composed of more skilled players tending to decide more quickly—but only until A reaches 81.2. Moreover, note the low R^2 in this regression, which indicates that neither group size nor ability explains much of the variation in lag times.

Next, we turn to *accuracy* rather than *speed*. For the regression in column 6, we define a new left-hand variable, *Correct*, which is

[19] This looks like the classic symptoms of extreme multicollinearity, but in fact the correlation between A (the group average) and *Best* is only 0.67. Replacing A—which, of course, includes *Best*—with the median does not reduce the multicollinearity at all (the correlation between the median and *Best* is also 0.67), and it generally produces worse-fitting regressions. For these reasons, we stick with the mean, rather than the median, in what follows.

[20] This is an admittedly narrow concept of heterogeneity. But, other than the sex composition of the group (which did not matter), it is the only measure of heterogeneity we have.

equal to 1 if the group's initial interest rate move is in the correct direction—i.e., if a rise in G is followed by a monetary tightening, or a decline in G is followed by a monetary easing—and equal to 0 otherwise. Do larger groups derive their advantage by being more accurate, in this sense?[21] Apparently not. The group-size dummy again shows no difference between groups of size four and size eight. It is interesting to note that, now, the average ability of the members of the group is also of no use in predicting the group's success—a surprising finding.

Having failed so far, we turn to one last performance metric: the frequency of interest rate changes. Remember that each change in the rate of interest costs the group a 10-point charge. So it is possible that larger groups do better because they "fiddle around" less with interest rates. To find out, we define a new left-hand variable, *Frequency*, which measures the number of rate changes a group makes over the course of a twenty-quarter game. Since interest rate changes are costly, it pays for groups to economize on them. The simple regression in column 7 reveals a modest effect of group interaction in producing more "patient" decision making. And, strikingly, the *Ability* variable seems to have little to do with the frequency of rate changes.

Here at last we find a partial answer to the question of why larger groups perform slightly better: they average 0.27 fewer interest rate changes per game (with a p-value of 0.08). Since only about 2.25 changes are made on average, this is a meaningful difference.

To summarize this investigation, larger groups take about as much time (measured in terms of data) and are about as accurate in their decisions as smaller groups. However, they make slightly fewer interest rate changes overall and, in this (limited) sense, are slightly more "stodgy" decision makers than individuals. This slightly more patient behavior, in turn, produces a systematic, though quite modest, performance advantage over small groups.

[21] Of course, since *Correct* is binary, a linear probability specification may not be appropriate. As an alternative, we could have performed a probit regression at the cost of not being able to cluster standard errors. The results from probit regressions are qualitatively and quantitatively similar to the linear probability specifications reported here.

Why might larger groups do slightly better? In this environment of pervasive uncertainty, each member of a group is likely to carry in a somewhat different notion of how the model works from his or her own experience with individual play—and thus, in particular, a different notion of how often to change interest rates. Group play allows members to pool the wisdom gained from their individual experiences. If pooling offers gains, but the gains are subject to diminishing returns, we might find groups of eight outperforming groups of four.

But then why are the gains from larger group size so small? One possibility might be that the optimal committee size is, say, $n = 6$. In that case, committees of four (too small) and eight (too large) might be almost equally suboptimal.[22] Alternatively, it could be that $n = 4$ and $n = 8$ are simply too close together, and that experimenting with, say, $n = 12$ or more might have produced larger differences. Finally, it is worth noting that our committees are all symmetric—everyone does exactly the same thing. By contrast, many real-world MPCs allow (or require) their members to specialize in certain tasks. Diminishing returns presumably sets in more slowly in such specialized committees than in symmetric committees.

5. Does Leadership Enhance Group Performance?

As noted in the introduction, virtually all decision-making groups in the real world, and certainly all MPCs, have well-defined leaders—e.g., the chairman of a committee. To an economist, or to a Darwinian evolutionist for that matter, this observation creates a strong presumption that leadership must be productive—for why else would it be so ubiquitous? But, as we show now, our experimental findings say otherwise: surprisingly, groups with designated leaders do *not* outperform groups without leaders.

We start table 4, as we did table 3, with a simple regression comparing the scores of groups with and without leaders—ignoring, for the moment, both average ability and group size. The leader dummy, defined to be 1 if the group has a designated leader and 0 otherwise, actually gets a *negative* (though insignificant) coefficient

[22]This possibility was suggested to us by Petra Geraats.

Table 4. Regression Results on Leadership

Dependent Variable→	**(1) Score**	**(2) Score**	**(3) Score**	**(4) Score**	**(5) *Lag***	**(6) *Correct***	**(7) *Frequency***	**(8) Score**
Leader Dummy	−0.832	−0.160			−0.287	−0.025	0.154	−0.718
	(1.225)	(0.742)			(0.415)	(0.033)	(0.152)	(1.098)
Ability		10.300***	12.257*	17.820***	−2.377***	0.009	−0.259	9.430***
		(3.515)	(6.098)	(4.138)	(0.822)	(0.102)	(0.347)	(3.182)
$Ability^2$		−0.064***	−0.078*	−0.114***	0.015**	0.000	0.002	−0.058***
		(0.023)	(0.041)	(0.028)	(0.006)	(0.001)	(0.002)	(0.021)
Best			−0.384					
			(2.697)					
$Best^2$			0.005					
			(0.017)					
Female				−1.164				
				(1.100)				
Large-Group Dummy								0.769
								(0.837)
Large Group × Leader								1.045
								(1.439)
Constant	87.054***	−325.405**	−393.587*	−609.677***	99.346***	0.352	10.582	−292.001**
	(0.613)	(133.642)	(202.219)	(153.455)	(30.251)	(3.825)	(13.029)	(120.957)
No. of Observations	504	504	264	252	504	504	504	504
R^2	0.01	0.23	0.32	0.32	0.07	0.01	0.02	0.03

Notes: Robust standard errors clustered by session are in parentheses. *, **, and *** denote significance at the 10 percent, 5 percent, and 1 percent level, respectively. "Leader Dummy" equals 1 if the group had a designated leader. "*Ability*" is the average score of all members of a group during part 1 of the experiment. "*Best*" is the highest average score attained by a member of a group during part 1. "*Female*" equals 1 if the leader was a female. "Large-Group Dummy" equals 1 if the group contained eight members.

in column 1. Once we control for ability in column 2, this small coefficient drops to virtually zero, and the effect of ability on group performance resembles that in table 3—a quadratic in A that peaks at 80.4. Thus the counterintuitive finding is that leadership does *not* affect group performance. We proceed now to try to overturn this surprising non-result.

One obvious explanation might be that our designated leaders achieve their high scores during part 1 purely by chance, and thus are really no more able than the others. This possibility is easily dismissed, however, by looking at scores in part 3 of the game—when subjects play again as individuals. Across all individuals who participated in the sessions with designated leaders, the correlation between their part 1 scores and their part 3 scores is 0.45, indicating substantial and durable individual effects. Thus it is not just luck; the leaders really are better.

One interesting question to ask, once again, is whether the group's score is driven more by the skill of the average member or by the skill of the leader. To address this question, we restrict our attention in column 3 to sessions with designated leaders (thus reducing the sample size to 264) and add the previously defined variables *Best* and $Best^2$ to the regression. Remember that *Best* is the average score of the highest-scoring individual in part 1—and thus the score of the designated the leader in part 2.

Interestingly, the average skill of the group's members ("*Ability*") is a much better predictor of performance than the skill of the leader ("*Best*"). To see this formally, we ran F-tests to determine the effect of omitting the two *Ability* variables versus omitting the two *Best* variables from the regression. For the *Ability* variables, the F-statistic is 8.7 ($p = 0.00$) whereas for the *Best* variables, the F-statistic is only 3.2 ($p = 0.06$). The comparative weakness of the *Best* variables illuminates the puzzling absence of leadership effects on performance: while the leader is the best player, he or she seems incapable of improving the performance of the group.[23]

We next ask whether leadership effects on group performance differ by the gender of the leader by adding a dummy variable *Female*

[23] The inverted quadratic in *Best* seen in column 3 looks peculiar, but it is upward sloping in the relevant range. Given the imprecision of the estimates of these coefficients, one shouldn't make much of this result.

to the regression in column 4. Again, we restrict our attention to sessions with designated leaders.[24] While the estimated coefficient for female leaders is negative, it does not come close to statistical significance. Thus, women do neither better nor worse as leaders.[25]

So leaders seem to have no discernible effect on their group's *score*. But do they influence the group's *strategy*? To examine this question, we look next at the dependent variable *Lag* defined earlier. Column 5 shows that leadership does not influence the speed of reaction significantly. While the coefficient of the leader dummy is negative, it is insignificant.

What about leadership effects on the likelihood of moving in the correct direction on the first interest rate change? The column 6 regression also shows essentially no effect.

Finally, we turn to the frequency of rate changes. Do groups with designated leaders change interest rates more (or less) frequently? The answer is (weakly) more frequently, as the regression in column 7 shows. But, again, the effect does not come close to statistical significance.

To this point, we have looked for leadership effects on the (tacit) assumption that they are the same in large ($n = 8$) and small ($n = 4$) groups. Similarly, in the previous section we examined the effects of group size while maintaining the hypothesis that size effects are the same with and without leaders. To test for possible interaction effects, the last regression in table 4 includes dummies for both group size and leadership, allowing an interaction between the two.

Column 8 reveals that the interaction effect is totally insignificant (p-value = 0.47). Still, if the positive coefficient is taken at face value, the regression suggests a small *negative* effect of leadership in smaller groups and a small *positive* effect in larger groups.

A fair summary so far would be to say that you need a magnifying glass—and you must ignore statistical significance—to see any effects of leadership on group performance. The main message, surprisingly, is that leadership does not seem to matter.

---

[24] A leader in one of the eight-person sessions refused to identify his or her gender, which reduced the number of observations to 252.

[25] They are also neither better nor worse as followers. The sex composition of the group does not help explain the group's performance.

Table 5. Improvements from Individual to Group Play, by Treatment

(1) Treatment	(2) Part 1 Mean Score (Individual Play)	(3) Part 2 Mean Score (Group Play)	(4) Difference
$n = 4$, no leader	78.4	87.1	8.7 (11.1%)
$n = 4$, leader	75.5	84.1	8.6 (11.4%)
$n = 8$, no leader	76.8	87.1	10.3 (13.4%)
$n = 8$, leader	78.2	88.4	10.2 (13.0%)

One other place to look for leadership effects is in how much people learn from their experience playing as a group. In our Princeton and Berkeley experiments, we found significant improvements in performance when individuals came together to play as groups. Could it be that the learning that takes place in group play is greater when the group has a designated leader?

Table 5 displays the *improvements* in score from part 1 (individual play) to part 2 (group play) separately for each of the four experimental treatments. Column 4 shows no support for the idea that group interactions help subjects more when there is a designated leader.

To assess statistical significance, we examine the dependent variable $DIFF_i$ suggested by table 5: the average score of a given subject in group play (part 2 of the game) *minus* that same individual's average score while playing as an individual in part 1. Table 5 suggests that improvements are slightly higher with larger groups but are independent of leadership. Thus, we include as right-hand variables dummies for both group size and whether the group is led or not. As usual, we cluster by session to obtain

$$\begin{array}{llll} DIFF_i = 8.71 & +\ 0.03\ LED_i & +\ 1.46\ D8_i & R^2 = 0.005 \quad N = 250 \\ \quad\quad (0.83) & \quad (0.99) & \quad (0.99) & \\ \quad\quad t = 10.5 & \quad t = 0.03 & \quad t = 1.5 & \end{array} \tag{7}$$

where LED is the leader dummy and $D8$ is the large-group dummy.

Table 6. Improvements from Part 1 to Part 3, by Treatment

(1) Treatment	(2) Part 1 Mean Score (Individual Play)	(3) Part 3 Mean Score (Individual Play)	(4) Difference
$n = 4$, no leader	78.4	83.2	4.8 (6.1%)
$n = 4$, leader	75.5	85.2	9.7 (12.8%)
$n = 8$, no leader	76.8	85.1	8.3 (10.8%)
$n = 8$, leader	78.2	84.9	8.7 (8.6%)

This regression shows that leadership has no effect on the improvement between individual and group play. On the other hand, participation in larger groups does improve upon individual performance slightly more than participation in smaller groups; however, the result does not quite rise to the level of statistical significance ($p = 0.15$).

One final question about leadership and learning can be raised. We found in both experiments that scores typically improve quite a bit when subjects move from individual play to group play (from part 1 to part 2) but then fall back somewhat when they return to individual play (from part 2 to part 3). The change in an individual's performance from part 1 to part 3 can therefore be used as an indicator of what might be called the "durable learning" that emerges from experience with group play. Is this learning greater when the group has a designated leader than when it does not?

Table 6 suggests that the answer is no. The subjects learn more from group play with a designated leader when $n = 4$, but less when $n = 8$. Notice, by the way, that the largest improvement in table 6 comes in the $\{n = 4, \text{leader}\}$ groups—the treatment that, by chance, got the weakest players.

The statistical significance of this result can be appraised by regressing the dependent variable $POSTDIFF_i$, defined as the difference between the average score of a given subject in part 3 of the game less that same individual's average score in part 1, on dummy variables for leadership and size. Clustering by session as usual, the result is

$$\begin{array}{llllll} POSTDIFF_i = 7.38 & +\ 0.41\,LED_i & -\ 0.18\,D8_i & R^2 = 0.00 & N = 250 & \\ \ (1.13) & \ (1.21) & \ (1.21) & & & \\ \ t = 6.5 & \ t = 0.3 & \ t = 0.2 & & & (8) \end{array}$$

This regression shows that neither group size nor leadership affects the durable performance gains that arise from exposure to group play.

In sum, there is no evidence from our experiment of superior (or even faster) performance by groups with designated leaders versus groups without. Overall, the most prudent conclusion appears to be that groups with designated leaders perform no differently than groups without leaders. This is a surprising finding, to say the least. Should we believe it? Maybe, but maybe not.

5.1 Why No Leadership Effects?

First, in defense of our experimental design, remember that we do *not* choose the leaders randomly or arbitrarily. Rather, each designated leader *earns* his or her position by superior performance *in the very task that the group will perform.* This principle for selecting leaders, we believe, imbues them with a certain legitimacy—as is normally the case in real-world groups. A second element of realism derives from the reward structure. Doubling the leader's reward in group play gives him or her a greater stake in the outcome—just as leaders of real-world groups normally have greater stakes in the outcome than other members do. For example, history will appraise the performance of the "Bernanke Fed" and the "Roberts Court." The names of most of the other members will be lost to history.

Second, however, while giving the leader the tie-breaking vote allows him or her to influence the group's decisions *in principle*, it may not do so *in practice*. For example, we found in Blinder and Morgan (2005) that there was no difference in either the quality or speed of group decision making when groups made decisions unanimously rather than by majority rule. And, as noted earlier, tie votes were rare in this experiment.

Third, and in a similar vein, we are able to test only for differences between groups with and without an *officially designated* leader; we have no independent measurement of how *effective* leadership is. Thus, some of our putative leaders may actually be quite passive, while strong leadership could emerge spontaneously in some of the groups without a designated leader.

Fourth, it should be noted that the task in our experimental setup is what psychologists call intellective (figuring something out) rather than, say, judgmental or moral (deciding what's right and wrong). So the surprising conclusion that leadership in groups has no apparent benefits should, at the very least, be limited to such intellective tasks. As Fiedler and Gibson (2001, 171) pointed out, "Extensive empirical evidence has shown that a leader's intellectual ability or experience does not guarantee good [group] performance." That said, making monetary policy decisions is, for the most part, an intellective task.

Fifth, however, there is never any disagreement among members of our ersatz MPCs over what the group's objectives are (including the relative weights). Every player tries to maximize exactly the same function. By contrast, there is potential for disagreement over the central bank's objectives and/or weights on at least some real-world MPCs (e.g., the FOMC)—which might allow more scope for effective leadership. In fact, this raises a broader issue. Our student subjects are arguably a more homogeneous group than at least some MPCs, to which people of diverse backgrounds are deliberately appointed.

Sixth, and related, our committees deal only with "normal" monetary policy decisions. It is certainly possible that greater scope for leadership might emerge if our experimental subjects were faced with crises, such as the ones the Federal Reserve and the ECB have been confronted with in 2007 and 2008.

Finally, and perhaps most important, our narrow experimental concept of leadership—leading the discussion, reporting the group's decision, and breaking a tie if necessary—does not correspond to the common meaning of "leadership" as expressed, e.g., in the admittedly chauvinistic statement "He's a leader of men." Our experimental leaders do not lead in the sense that a military officer leads a platoon, a politician leads a party, or an executive leads a business. Brown, Scott, and Lewis (2004) classified leaders as "transformational" and "transactional," the latter meaning motivating subordinates with rewards. Our experimental leaders were neither.

We thought about trying to select group leaders by what might loosely be described as "leadership qualities" but quickly abandoned the idea as being too subjective and too difficult. We think this decision was the right one. But, in interpreting the experimental results, it is important to remember that our leaders are selected, on average, for their "smarts," not for their "leadership qualities." There is no reason to think that the cognitive ability we use to select group leaders correlates highly with the traits that are associated with leadership in the real world, such as verbal dexterity, aggressiveness, an extroverted personality, a trustworthy affect, good looks, and height. However, we certainly hope (and believe) that cognitive ability is a relevant consideration in the selection of real-world central bank heads.

Similarly, it seems plausible that true—as opposed to putative—leadership in groups may need to emerge slowly over time, as the leader demonstrates good performance and as other members grow to respect his or her judgment, acumen, and group-management skills. A one-time, ninety-minute laboratory experiment leaves no scope for that sort of leadership to emerge.

Thus we certainly do not believe that our experimental results provide the last word on leadership effects. We offer them as something closer to the first word, and invite other researchers to pick up the challenge.

6. Conclusions

In this paper, we replicate earlier findings from Blinder and Morgan (2005) showing that simulated monetary policy committees make systematically better decisions than the same individuals making decisions on their own, without taking any longer to do so. This experimental evidence supports the observed worldwide trend toward making monetary policy decisions by committees, rather than by lone-wolf central bankers. We also find several suggestive shreds of evidence that the margin of superiority of groups over individuals is greater when the individuals are of lower ability.

But the more novel findings of this paper pertain to groups that differ in terms of size and leadership. We find some weak evidence that larger groups (in our case, $n = 8$) outperform smaller groups ($n = 4$), mainly because larger groups seem better able to resist the temptation to "fiddle" with interest rates too much. But these differences are small, and many are not statistically significant. So, in terms of institutional design, it is not clear whether larger or smaller MPCs are to be recommended. (Remember that $n = 7$ is the mean and modal size of real-world MPCs.)

Our most surprising and important result, at least to us, is that ersatz MPCs do *not* perform any better when they have a designated leader than when they do not—even though every real-world MPC has a clear (and sometimes dominant) leader, and even though our designated leaders were chosen on the basis of their skill in making monetary policy. We caution that we would not apply this finding beyond the realm of intellective tasks—e.g., we do not recommend that army platoons venture out without a commanding officer!

But that said, there are probably many more intellective than combative tasks in the economic world, certainly including monetary policy. For example, promotions in business are often based on superior performance on metrics that are basically intellective. So this finding, if verified by other work, is potentially of wide applicability. In terms of the taxonomy of MPCs emphasized by Blinder (2004), our results suggest that an *individualistic* committee, where the leader is only modestly more important than the other members, may function just as well as a *collegial* committee, where the role of the leader is more pronounced.

Appendix. The Instructions

Note: Portions of the instructions read only during sessions with leaders are enclosed in brackets.

In this experiment, you make decisions on monetary policy for a simulated economy, much like the Federal Reserve does for the United States. At first, you will make the decisions on your own; later, we will bring you all together to make decisions as a group. [At that point, one of you will be designated the leader of the group, as I will explain shortly.]

We have programmed into each computer a simple model economy that generates values of unemployment and inflation, period by period, for twenty periods. Think of each period as a calendar quarter, so the game represents five years. Each quarterly value of unemployment and inflation depends on the interest rates you choose and on some random influences that are beyond your control. Every machine has exactly the same model of the economy, but each of you will get different random drawings and so will have different experiences.

Your goal is to keep unemployment as close to 5 percent, and inflation as close to 2 percent, as you can—quarter by quarter. As you can see from the top line on the screen, we start you off with an interest rate of 7 percent in period 1. Initially, unemployment and inflation differ slightly from the targets of 5 percent and 2 percent because of the random influences I just mentioned. Starting with period 2, you must choose the interest rate.

Raising the interest rate will *increase* unemployment and *decrease* inflation. But the effects are delayed—neither unemployment nor inflation responds immediately. Similarly, lowering the interest rate will decrease unemployment and increase inflation. But, once again, the effects are delayed.

The computer determines your score for each period as follows. Hitting 5 percent unemployment and 2 percent inflation exactly earns you a perfect score of 100 points. For each tenth-of-a-point by which you miss each target, you lose a point on your score. Direction doesn't matter; you lose the same amount for being too high as for being too low. Thus, for example, 5.8 percent unemployment and 1.5 percent inflation will net you a score of 100 *minus* 8 points for missing the unemployment target by eight-tenths *minus* 5 points for missing the inflation target by five-tenths, or 87 points. Similarly, 3.5 percent unemployment and 3 percent inflation will net you 100 – 15 – 10, or 75 points. If you look at the top line of the display, you can see that the initial unemployment rate of 5 percent and inflation rate of 1.9 percent yields a score of 99.

Finally, there is a cost of 10 points each time you change the interest rate. The 10 points will be deducted from that period's score.

Are there any questions about the scoring system?

As you progress through the experiment, accumulating points both in individual and in group play, the computer will keep track of your cumulative *average* score on the 1–100 scale. At the end of the session, your cumulative average score will be translated into money at the rate of *25¢ per point*, and you will be paid your winnings by check. Thus a theoretical perfect score of 100 would net you $25, an average score of 80 would give you 80 percent of $25, or $20, etc. You are guaranteed a minimum of $15, no matter how badly you do.

[When you play as individuals, everyone is treated the same. But when we bring you together to play as a group, one of you will serve as the group's *leader*. The leader will be the one who scored the highest in individual play, and he or she will receive *twice* as many points during group play.]

The game works as follows. You can move the interest rate up or down, in increments of 1 percentage point, by moving the slide bar on the left-hand side of the screen, or by clicking on the up or down buttons. Try that now to see how it works. When you have selected the interest rate you want, click on the button marked "Click to Set Rate." Do that now to see how it works. The computer has recorded your choice, drawn the random numbers I mentioned earlier, and calculated that period's unemployment, inflation, and score.

There is one final, important aspect to the game. At a time period selected at random, but equally likely to be any of the first ten periods, aggregate demand will *either* increase *or* decrease. You will not be told *when* this happens nor in *which direction.*

If aggregate demand *increases*, that tends to push unemployment down and, with a lag, inflation up. If aggregate demand *decreases*, that tends to push unemployment up and, with a lag, inflation down. The essence of your job is to figure out when and how to adjust interest rates in order to keep unemployment as close to 5 percent, and inflation as close to 2 percent, as possible.

Remember, the change in aggregate demand comes at a randomly selected time within the first ten periods, and we will not tell you whether demand has gone up or down. And each interest rate change will cost you 10 points in the period you make it.

Are there any questions?

Please sign the consent form located in the folder next to your computer.

This will all be simpler once you've practiced on the apparatus a bit. You can do so now, and your scores will just be displayed for your information; they will not be recorded or counted. You can practice for about five minutes to develop some familiarity with how the game works. During this practice time, feel free to ask any questions you wish.

OK, it's time to start the game for real now.

In the first part of the experiment, you will play the monetary policy game twelve times *by yourselves*. After you have played the game twelve times, the computer will prevent you from going on. You may not communicate with any other player, and the points you earn will be your own.

[As I mentioned, the student who earns the highest score in this part of the experiment will be the leader of the group in the next part.]

Please start now by clicking the continue button, and proceed at your own pace.

Now please gather around the projection screen to play the same game as a group. [The highest scoring player will be the leader.]

In this part of the experiment, you will play exactly the same game twelve times. The rules are the same except that decisions are now made *by majority rule*. [In case of a tie, the leader will cast the tie-breaking vote. The leader will control the mouse.] You may communicate freely with each other, as much as and in any way you wish. While playing as a group, each of you will receive the group's score [except the leader, who will earn twice as many points]. At the conclusion of group play, we will show you how your performance compared with the top scores achieved when we ran this game at Princeton. Any questions?

Please begin.

OK. Now please return to your individual seats and, once again, play twelve rounds of the game by yourselves. Communication with other players is not allowed. The computer will again stop you after twelve rounds.

Please begin.

OK. The experiment is now over. Thank you for participating.

References

Blades, J. W. 1973. "Influence of Intelligence." In *The Influence of Intelligence, Task Ability, and Motivation on Group Performance*, ed. J. W. Blades and F. E. Fiedler, 76–78. University of Washington, Seattle: Organizational Research Technical Report.

Blinder, A. S. 2004. *The Quiet Revolution: Central Banking Goes Modern.* New Haven, CT: Yale University Press.

Blinder, A. S., and J. Morgan. 2005. "Are Two Heads Better Than One? Monetary Policy by Committee." *Journal of Money, Credit, and Banking* 37 (5): 789–812.

Brown, D., K. Scott, and H. Lewis. 2004. "Information Processing and Leadership." In *The Nature of Leadership*, ed. R. Sternberg et al. New York, NY: Sage Publications.

Chappell, H. W., Jr., R. R. McGregor, and T. Vermilyea. 2005. *Committee Decisions on Monetary Policy.* Cambridge, MA: MIT Press.

Edmondson, A. 1999. "Psychological Safety and Learning Behavior in Work Teams." *Administrative Science Quarterly* 44 (4): 350–83.

Erhart, S., and J. L. Vasquez-Paz. 2007. "Optimal Monetary Policy Committee Size: Theory and Cross-Country Evidence." Paper presented at Norges Bank conference.

Fiedler, F. E., and F. W. Gibson. 2001. "Determinants of Effective Utilization of Leader Abilities." In *Concepts for Air Force Leadership*, ed. R. I. Lester and A. G. Morton, 171–76. Maxwell Air Force Base, AL: Air University Press.

Guth, W., M. Vittoria Levati, M. Sutter, and E. van der Heijden. 2004. "Leadership and Cooperation in Public Goods Experiments." Discussion Paper on Strategic Interaction No. 2004-29, Max Planck Institute of Economics.

Lombardelli, C., J. Proudman, and J. Talbot. 2005. "Committees versus Individuals: An Experimental Analysis of Monetary Policy Decision Making." *International Journal of Central Banking* 1 (1): 181–205.

Mahadeva, L., and G. Sterne, ed. 2000. *Monetary Policy Frameworks in a Global Context*. New York, NY: Routledge Publishers.

Maier, N. R. F. 1970. *Problem Solving and Creativity in Individuals and Groups*. Belmont, CA: Brooks/Cole.

Rudebusch, G. 2001. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." *Review of Economics and Statistics* 83 (2): 203–17.

Sibert, A. 2006. "Central Banking by Committee." *International Finance* 9 (2): 145–68.

White, H. 1980. "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity." *Econometrica* 48 (4): 817–38.

Inflation Targeting and Target Instability*

Robert J. Tetlow
Board of Governors of the Federal Reserve System

Monetary policy is modeled as being governed by a known rule, except for a time-varying target rate of inflation. The variable target can be thought of either as standing in for discretionary deviations from the rule or as the outcome of a policymaking committee that is unable to arrive at a consensus. Stochastic simulations of FRB/US, the Board of Governors' large rational-expectations model of the U.S. economy, are used to examine the benefits of reducing the variability in the target rate of inflation. We find that putting credible boundaries on target variability introduces an important nonlinearity in expectations. The effect of this is to improve policy performance by focusing agents' expectations on policy objectives. But improvements are limited; it does not generally pay to reduce target variability to zero. More important, this nonlinearity in expectations allows for policy to be conducted, at the margin, with greater attention to output stabilization than would otherwise be the case. The results provide insights as to why inflation-targeting countries use bands and why the bands they use are narrower than studies suggest they should be. A side benefit of the paper is the demonstration of a numerical technique that approximates to arbitrary precision a nonlinear process with a linear method, thereby greatly speeding and making more robust the computation of simulation results.

JEL Codes: E3, E5, C6.

*The views expressed in this paper should not be attributed to the Board of Governors, the FOMC, or staff members of the Federal Reserve System. I am thankful for the guidance and suggestions of the editor, John Taylor, and two anonymous referees. In addition, Steve Ferris, Ben Hunt, Tiff Macklem, Athanasios Orphanides, Dave Reifschneider, Larry Schembri, John C. Williams, and seminar participants at Carleton University, the Bank of Canada, and the Reserve Bank of Australia provided useful comments. All remaining errors are my own. E-mail: Robert.J.Tetlow@frb.gov. Tel: (202) 452-2437. Fax: (202) 452-5296. Web page: www.roberttetlow.com.

1. Introduction

> While the pursuit of price stability does not rule out misfortune, it lowers its probability. . . . If households are convinced of price stability, they will not see variations in relative prices as reasons to change their long-run inflation expectations. Thus, continuing to make progress toward this legislated objective will make future supply shocks less likely and our nation's economy less vulnerable to those that occur.
>
> (Greenspan 1998)

In recent years, a number of countries have adopted inflation targeting as the legislated objective for their countries, in pursuit of the same bounty that Alan Greenspan, then Chairman of the Board of Governors of the Federal Reserve System, outlines in the above quotation. New Zealand (1990), Canada (1991), the United Kingdom (1992), Finland (1993), and Sweden (1993), among other countries, have established explicit targets for inflation. The Federal Reserve, on the other hand, makes broad statements about "price stability" being the ultimate objective of policy, without defining quantitatively what the term means. Is this sufficient? What about announcing a "comfort zone," within which any rate of inflation is as good as any other, as in the case of Australia?[1]

As Chairman Greenspan's statement emphasizes, a key purpose of public declarations regarding policy objectives is to direct private agents' expectations toward policy objectives. Doing so successfully builds up reputation for policy, and policymakers, making the economy "less vulnerable" to adverse shocks. In this paper, we examine what role a limited form of policy credibility—specifically, the idea that *perceived inflation targets* might drift within bounds over time—might have on monetary policy outcomes.

[1] Siklos (1997) provides a nice short summary of the practices of various inflation-targeting countries; see especially the table on pp. 132–33.

We model perceived inflation targets as a bounded random walk with a band width that is under the control of the Federal Reserve. We think of this band width as representing either the degree of discretion that the Federal Open Market Committee (FOMC) allows itself or the byproduct of inertial responses by a committee that cannot come to consensus on the target over time. We embed this process for the target within the rational-expectations version of the Federal Reserve Board's macroeconometric model of the U.S. economy, FRB/US, and conduct stochastic simulations to provide a quantitative guide to the possible benefits that might be accrued with a regime where there is randomness in policy but only within bounds. We show that target bands for inflation control imply an important nonlinearity that increases in significance as the target bands narrow. A Federal Reserve that chooses the target band would find that narrow bands are beneficial for the economy because of their effect on providing a focal point for private agents' expectations. One contribution of the paper is the resolution of the numerical problem presented by the nonlinearity in combination with the large scale of the model.

To presage the results, we find that constraining the perceived drift in inflation targets engenders a "honeymoon effect" that facilitates inflation control above and beyond what arises from policy actions alone. More generally, for a given policy rule, narrowing the bands improves economic outcomes, although the gains decline as the bands get narrower. More important is that greater credibility of monetary policy in the form of a narrower perceived range of the target rate allows the central bank to be less hawkish in setting policy rates than otherwise would be the case. We argue that, together, these results favor point targets for inflation over target ranges, or "comfort zones."

The rest of the paper proceeds as follows. In section 2 we discuss variability in monetary policy, and how the behavior of central banks in general and the Federal Reserve in particular might be reasonably characterized by drifting but bounded targets for inflation. We also describe how "variability" is operationalized for this paper. The third section provides a brief encapsulation of the FRB/US model. Section 4 then discusses aspects of our simulation technique. The fifth section summarizes our results. A sixth section sums up and concludes.

2. Inflation Targeting and Variable Targets

2.1 Policy Rules and Policy Objectives

A formal way of characterizing monetary policy is the solution to a dynamic optimization problem. Another quite different method is where, in the words of Blinder (1998), decision makers "look out the window" and adjust policy according to whether the economy is currently running "too hot" or "too cold." The former method is an example of policy by rules; the latter is inherently discretionary. When the Federal Reserve, in pursuing discretionary policy, failed in the 1970s to adjust the funds rate sufficiently in response to inflationary shocks, the implication was that the implicit inflation target was being increased—not necessarily as a deliberate act of policy but perhaps as an implication of the Federal Reserve's unwillingness to forgo output objectives in order to maintain a fixed inflation target.[2] Between these two extreme characterizations of policy lie a wide range of alternatives embodying differing levels of commitment and flexibility.

In this paper, we hypothesize that the public has information about the broad objectives of policy but no information regarding a specific target. To illustrate, suppose Federal Reserve watchers were to "invert" a policy rule, like the Taylor (1993) rule, on the target rate of inflation, π_t^*:[3]

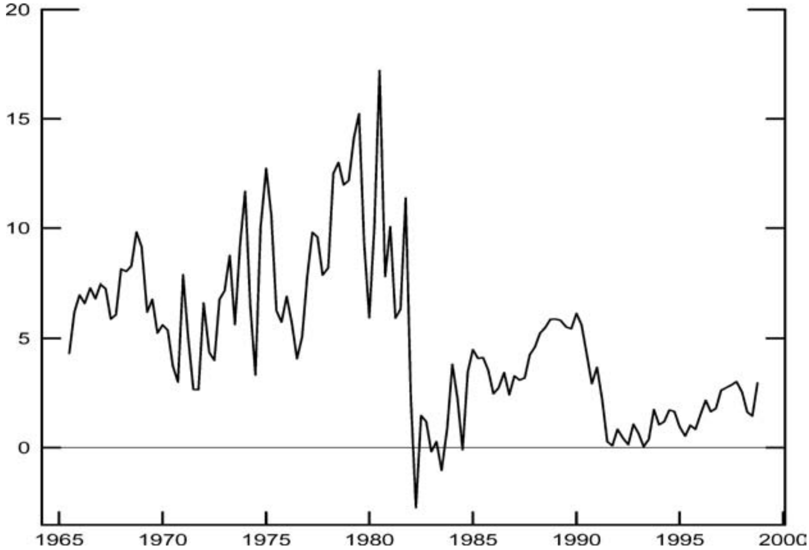
$$R_t = 2 + \tilde{\pi}_t + 0.5\left(\tilde{\pi}_t - \pi_t^*\right) + 0.5y_t, \tag{1}$$

where rule R is the nominal federal funds rate; the equilibrium real rate, rr^*, is taken, as Taylor did, to be 2 percent as shown; $\tilde{\pi}$ is the

[2]This characterization of policy is consistent with the rhetoric of Meltzer (1991, 34) when he accuses the Federal Reserve of having lost control of inflation in the 1970s because FOMC "members were reluctant to allow interest rates to rise by as much as required by staff forecasts. They hoped to achieve lower inflation by reducing money growth but were reluctant to allow interest rates rise by relatively large steps."

[3]A computer search of speeches by Board members over the period from 1993 to 1999 for the words "Taylor rule" and derivatives thereof reveals twelve hits for speeches by then Governors Meyer, Gramlich, Ferguson, and Yellen. No other rule had ever been publicly cited in speeches. The lessons of the Taylor rule have been reported to members of the Board "as principles to guide decisions about the setting of the federal funds rate at FOMC meetings" (Meyer 1998, 16).

Figure 1. Target Rate of Inflation as Implied by the Taylor Rule



four-quarter change in the chain-weight PCE price index; and y is the output gap.[4] An asterisk indicates an equilibrium or target level, as applicable. This results in the time series shown in figure 1.

The figure shows, first, that there has clearly been a large shift in either the equilibrium real interest rate or the target rate of inflation, or both, in the early 1980s.[5] This helps explain Taylor's focus on the 1987–92 period as one where the fit is "remarkable." Second, over both subsamples there have been large and persistent movements in the implicit target. High-frequency movements in π^* could

[4]We replace the original Taylor article's GDP price deflator as the relevant inflation measure to reflect current FOMC emphasis; otherwise, our formulation is the same as the original, up to the time-varying target. Other rules would give different answers, but no parsimonious rule would significantly alter our basic point.

[5]It is likely that both a shift in π^* and a shift in rr^* occurred. The early 1980s was the period of the Volcker disinflation, in which CPI inflation was deliberately brought down by the Federal Reserve from about 12 percent to approximately 5 percent. Bomfim (1997) shows that the equilibrium real interest rate apparently rose at about the same time, perhaps due to the accumulation in the early 1980s of large public-sector deficits.

be attributed to a missing term from the rule, associated perhaps with tactical concerns in the conduct of policy, or control errors. But movements as persistent as those shown are difficult to ascribe to something other than changes in the target. Third, with the exception of the once-and-for-all shift at around 1980, the shifts in the implicit target, however persistent, appear to have been bounded within ranges.

The Federal Reserve has not made a regular habit of explicitly discussing the quantitative objectives of policy and so it seems reasonable to characterize the inflation target as a latent variable of policy—something that falls out of funds rate settings given the stochastic shocks that are borne by the economy—rather than the other way around. The "zone-quadratic" preferences of Orphanides and Wieland (2000) are consistent with this characterization; Orphanides and Wilcox (2002) describe something similar to this on "opportunistic disinflation" as a description of Federal Reserve behavior in the 1990s.[6] The rhetoric of late of the FOMC to the effect that the Committee has a "comfort zone" within which the bulk of the membership is satisfied with inflation is very much like this depiction.[7]

Quantitative work by Aksoy et al. (2006) provided both an assessment of opportunistic disinflation and a symmetric extension of the concept that permits drift, within bounds, of actual inflation, while the Federal Reseve concerns itself with output stabilization and then strong inflation control at the edges of the band. This describes a "zone of indifference" regarding inflation—not unlike the policy practiced by Australia—and, as Riboni and Ruge-Murcia (2008) show, can arise from committee structures that can agree on a point when inflation is "too high" or "too low" but cannot come

[6]It is noteworthy that opportunistic disinflation was the name that Federal Reserve Governor Laurence Meyer gave to the policy he saw the FOMC following when he joined the Board in early 1995. The Orphanides and Wilcox (2002) paper was written as a theoretical explanation of Meyer's observation. See Meyer (1997) for a description of opportunistic disinflation, including the author's claim to paternity.

[7]The term "comfort zone" appeared in the headline of a September 2002 *New York Times* interview of former Governor Laurence Meyer. Since then, the term has been used repeatedly by various FOMC members, including Bernanke (2005) and Yellen (2006). It is normally, and imprecisely, taken as a range between 1 and 2 percent for the four-quarter growth rate of either the headline or core PCE price index.

to a consensus on intermediate cases. Under this interpretation, the upper and lower boundaries for the target rate of inflation are points at which a critical mass on the committee would say that trend inflation is out of line with long-term objectives. In between, there might be more of a bias toward going back to the long-run target—in the sense that the probabilities point in that direction—but not an overwhelming impetus; some committee members might be inclined to await further information (and more shocks) before acting.[8] At the midpoint of the bands, all FOMC members are in agreement that inflation is not a problem.

The next subsection formalizes the idea of inflation targets as a bounded, random process.

2.2 Modeling Inflation Target Bands

The preceding subsection has established at least a prima facie case for treating the target of monetary policy as a random variable; in this subsection, we operationalize this idea. We begin by assuming that monetary policy can be described by a simple interest rate reaction function, as follows:

$$R_t = rr^* + \sum_{i=0}^{3} \pi_{t-i}/4 + 2.2 \left[\sum_{i=0}^{11} \pi_{t-i}/12 - \pi_t^* \right] + 1.5 y_t. \qquad (2)$$

Equation (2) differs from the canonical Taylor rule in three ways. First, the rule's inflation term is expressed as a twelve-quarter moving average of inflation instead of four quarters. Williams (2003) shows that, in the context of FRB/US, equations with longer lags on inflation outperform shorter lags. Second, the coefficients on both inflation and the output gap are larger for our rule. In fact, Williams (2003) shows that this parameterization is an efficient one in that no other coefficient values can produce lower unconditional variances of output and inflation simultaneously without increasing the

[8]The practice of decision making by consensus at the FOMC is well known. See, e.g., Blinder (1998) and Meyer (1998). Blinder notes that this can lead to sluggish responses. As we shall argue below, it may also lead to a disconnection between long-term objectives and short-term practices of the Federal Reserve.

variability of the federal funds rate.[9] Formally, the parameterization of equation (2) is the solution to a problem that minimizes the following loss function:

$$E_0 \sum_{i=1}^{\infty} \varphi \left[\tilde{\pi}_{t+i} - \pi^*_{t+i}\right]^2 + (1-\varphi) y^2_{t+i} \tag{3}$$

subject to the model, the specification of the rule equation (2), and to a constraint on maximum variability of the federal funds rate.[10] By choosing a (constrained) optimal rule in this way, we avoid specious results that might arise simply from the use of a poorly performing rule. In addition, we have the advantage of being able to consider policymakers' preferences directly. The parameterization of equation (2) is optimal for an authority with relatively "hawkish" inflation preferences, $\varphi = 0.75$; we shall examine alternative preferences a bit later.

The third difference from the Taylor rule is that the target rate of inflation, π^*_t, carries a time subscript; we allow time variation in the monetary policy target. We also assume that private agents know the rule but do not know the target on a period-by-period basis, so that they can infer the target only with a one-quarter lag. We consider two possible data-generating processes for the target rate of inflation: a random walk and a bounded random walk. (Clearly, no authority would truly permit the target rate of inflation to vary from plus to minus infinity, so the pure random-walk case should be thought of as a benchmark.) In both cases, the law of motion for the target looks the same:

$$\pi^*_t = \pi^*_{t-1} + \mu_t. \tag{4}$$

[9] Equation (2) also dominates rules with shorter moving averages in the inflation rate as judged by the same criterion of minimized unconditional variances. Note that, for expositional simplicity, we have arbitrarily rounded the many decimal places used in the rule actually used for simulation.

[10] The practice, following Williams (2003), among many others, of imposing a penalty on funds rate variability provides two benefits: (i) the optimally computed coefficients on the policy rule are of a magnitude that would be considered reasonable, and (ii) the implied variability of the federal funds rate is close to the historical experience. The same qualitative results obtain when funds rate volatility is mildly penalized instead of the hard constraint we use here.

The difference is in the distribution of innovations. In the pure random-walk case, the innovations, μ_t, are independently and identically distributed:

$$\mu_t \sim N(0, \sigma^2). \tag{5}$$

In the bounded random-walk case, innovations are bounded by the target bands. Letting $\bar{\pi}^*$ and $\underline{\pi}^*$ designate the upper and lower bounds for π^*, Johnson, Kotz, and Balakrishnan (1994, 156) show that the (truncated) probability density function of innovations is

$$\mu_t \sim \frac{1}{\sigma\sqrt{2\Pi}} \exp\left[\frac{-\mu^2}{2\sigma^2}\right] \left\{\frac{1}{\sigma\sqrt{2\Pi}} \int_{\left(\underline{\pi}^* - \pi^*_{t-1}\right)}^{\left(\bar{\pi}^* - \pi^*_{t-1}\right)} \exp\left[-\frac{\left(t - \pi^*_{t-1}\right)^2}{2\sigma^2}\right] dt\right\}^{-1}. \tag{6}$$

As ugly as equation (6) looks, it is simply a truncated normal distribution, with truncation points determined by band widths and the inherited rate of target inflation, given σ. That is, the truncation points, $\{\underline{\pi}^* - \pi^*_{t-1}, \bar{\pi}^* - \pi^*_{t-1}\}$, shift around with the inherited target inflation rate. We assume that the bands are spaced symmetrically around a given midpoint, which we set to zero.[11] Let us also assume that the bands on the target rate of inflation are $+1$ and -1 percent, respectively, as in our base-case simulations below. When the inherited target inflation rate happens to be zero, the distribution of innovations is simply the normal distribution with the tails beyond $+1$ and -1 chopped off. Besides being symmetric, with our base-case value of σ of 0.25, this distribution is functionally identical to the normal distribution because the truncation points are four standard deviations away. However, when the inherited target rate of inflation is positive, the upper bound on innovations moves to the right, eliminating more of the positive innovations that would otherwise be possible and adding more of the negative innovations. If the inherited target rate of inflation were exactly 1 percent, the

[11]We elect to ignore the zero lower bound on nominal interest rates in order to focus on the issue of interest. Other than the truncation described in the main text, the model is linear, so one can reinterpret the target rate as being at a rate where the likelihood of the zero-bound problem being binding is infinitesimally small, if so desired.

distribution of possible innovations would be exactly one-half of the normal distribution.[12]

The case of a fixed target rate of inflation is nested within equation (6), as $\underline{\pi}^* = \bar{\pi}^* = 0$. In addition, permitting $\bar{\pi}^* \to \infty$ and $\underline{\pi}^* \to -\infty$ gives the pure random-walk case.

It is important to note that these boundaries for the target do not imply the same boundaries for inflation itself. In the short run, many forces are at work on inflation—of which only one is monetary policy—and, in any case, inflation stabilization is not the sole concern of monetary policy. Thus, in general, inflation itself will be more variable than the target.

Two issues regarding the law of motion for π^* warrant discussion. First, we assume that π^* moves *randomly* within the target bands. As discussed above, the Volcker disinflation excepted, time variation in the de facto target of the Federal Reserve does not appear to have arisen from a conscious selection of different priorities. Instead, it appears to have come from an unwillingness to respond to shocks with sufficient strength to be consistent with a constant target. Indeed, the tendency in that period was, in the words of the Federal Reserve's then Director of Monetary Affairs (and current Board Vice Chairman) Donald Kohn, to "adjust nominal rates too slowly in the initial stages of a cycle, and then to overstay a policy stance" while awaiting further information.[13] While the process that Kohn alludes to will produce a correlation of shocks, this is only to the extent of the sign of the shocks—upward in the 1960s and 1970s, and downward in the 1990s.[14] In this context, the assumption adopted here is a tractable one and does not seem like

[12]We are modeling a discrete-time version of what is sometimes called reflecting barriers. With reflecting barriers, that proportion of the distribution of targets that would have called for targets outside of the bands is discarded. An alternative is absorbing barriers in which the truncated portion of the distribution is assigned to the boundary value, causing a tendency for target rates of inflation to mass at the boundary edges. With reflecting barriers, the target rate of inflation almost never reaches the boundary, as we shall see. Dixit (1993) provides a thorough discussion of the mathematics and modeling of both types of barriers in continuous time.

[13]See Kohn (1991, 101). As Meyer (1998, 18–19) notes, it is the Director of Monetary Affairs that presents the policy options to the FOMC.

[14]We might also note that to the extent to which innovations to the target were *systematically* correlated with shocks or other state variables, it should show up in the specification of the rule itself. And to the extent that these shocks are

an egregious oversimplification.[15] Second, we also assume that the bounds on the target are regarded as fully credible by private agents. Together, these two assumptions will permit us to model expected future inflation targets independent of the rest of the model.

2.3 Expected Future Inflation Targets

In the preceding subsection, the target rate of inflation was modeled with a truncated normal distribution. The truncation at any given date is determined entirely by (i) the inherited target inflation rate, (ii) the variance of innovations, and (iii) the band width. In this subsection, we describe the expected future target rate of inflation conditional on some initial value. The nature of this expectation is key to the results that follow because in the FRB/US model, expected *future* target rates of inflation feed into *current* price- and wage-setting behavior.

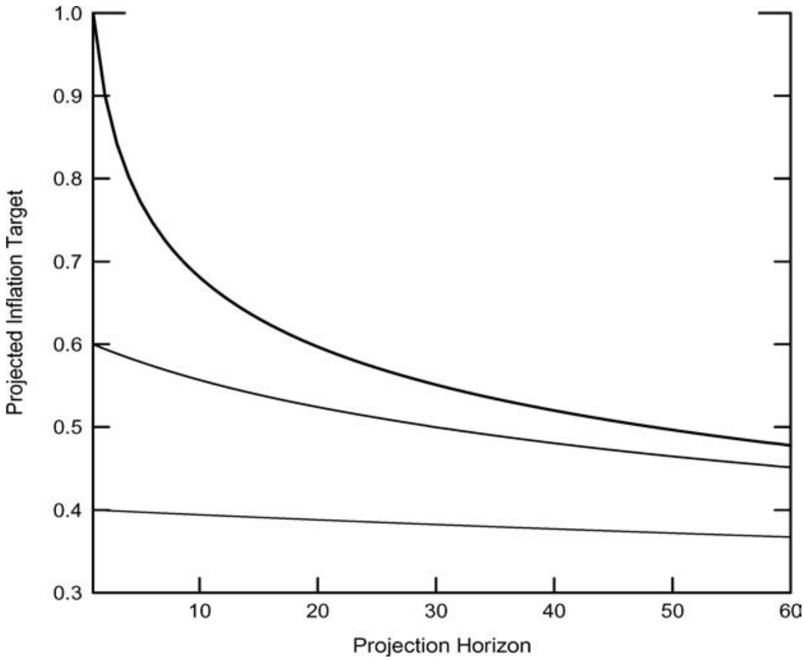
As noted, agents are assumed to know the process for the target rate of inflation, but not the current value of innovations. Since they know the inherited target inflation rate, the expected value of the current target can be computed, $E_{t-1}\pi_t^*$. And because innovations are independent, this value can then be taken as given and a value for $E_{t-1}\pi_{t+1}^*$ can be computed, and so on, in a straightforward application of the chain rule of forecasting.

At the midpoint of the bands the truncation is symmetric, the truncation points being equally distant from the inherited target rate; the pure random-walk prediction obtains in this special case. When $\pi^* = \bar{\pi}^* = 1$, the entire upper half of the normal distribution is truncated and only the lower half is pertinent. It follows that in this case the complete distribution of future target inflation rates is (weakly) less than the current target, and so the optimal prediction is for a decline in the future target and, symmetrically, for $\pi^* = -1$. More generally, it is clear that (i) the expected future target approaches the midpoint of the bands as time approaches infinity, and (ii) the initial rate of reversion toward the midpoint

correlated with the output gap, they may already be captured in the Taylor-rule coefficient for that variable.

[15] We note as well that this specification for the target rate of inflation allows a straightforward comparison of our results with those of the exchange-rate target-band literature that is related to the modeling undertaken here.

Figure 2. Agents' Inflation-Target Forecasts (Bounded Random Walk, Width = ±1)



will vary positively with the proximity of the inherited target rate to the bound, given the standard deviation of innovations. Johnson, Kotz, and Balakrishnan (1994, 156–62) show how to compute the probability density function for the truncated normal. Based on this, figure 2 shows the expected future path of $E_{t-1}\pi^*_{t+j}$ for three initial values of π^*_{t-1}, given $\sigma = 0.25$, and symmetric bands with the upper value arbitrarily set at $\bar{\pi}^* = 1$.[16]

The expected future path of target rates of inflation depends only on π^*_{t-1}, $\bar{\pi}^*$, and σ. Larger values of σ produce stronger nonlinearities because of higher probabilities of truncation for any given inherited target rate (other than zero). Symmetrically increasing the width of the target band, $\bar{\pi}^* - \underline{\pi}^*$, has the opposite effect.

[16]The standard deviation of innovations of 0.25 percent are taken from the standard deviation of changes in the Philadelphia survey of inflation expectations, ten years ahead, our proxy for the public's expectation of inflation in the long term.

Two other issues merit some attention. First, the discussion above has shown how the existence of credible bands serves to direct expectations of the target rate of inflation toward a fixed point as the horizon grows; the precise characteristics of the bands, together with the inherited target rate, determine the rate at which those expectations converge. This has the same flavor as the monetary policy frameworks employed by some central banks, most notably the Bank of England. In these frameworks, the central bank promises to bring inflation back to its target level *within a fixed horizon*. Just as in this paper, this introduces a nonlinearity into what might otherwise be a linear system and for the same purpose: to attempt to establish credibility. Second, for our base-case calibration, we have chosen $\sigma = 0.25$ on the basis of survey data for long-term inflation expectations. If one were to require learning on the part of private agents, rather than agents knowing the target with a one-period lag, one could then imagine that smaller shifts in the de facto target would support the same value of σ owing to the magnification that learning would entail.[17]

3. The FRB/US Model

In this section we outline the basic features of the FRB/US model. Limitations of space prevent us from doing justice to the features of the FRB/US model here; accordingly, we focus on those aspects that are germane to the issues studied in this paper. Our goal is to convince readers that the model is sensible and worthy of using to address the question at hand. Those interested in more information about the model can consult the working paper version of this article, as well as other publications.[18]

FRB/US is a relatively large rational-expectations structural model consisting of about thirty key behavioral equations and

[17]There would, of course, be other implications as well, most notably to the persistence of perceived shocks. We thank an anonymous referee for pointing out this angle in inquiry.

[18]Interested readers should consult Brayton and Tinsley (1996) on general aspects of the model. Particularly accessible descriptions of the modeling of expectations, and the monetary policy transmission mechanism, respectively, are Brayton et al. (1997) and Reifschneider, Tetlow, and Williams (1999). An application of the model to the analysis of disinflations is Bomfim et al. (1997).

several hundred identities. Rigidities in wage and price setting imply that monetary policy has effects on the real economy in the short run, but in the long run, policy can affect only the inflation rate.

The model is designed around the joint determination of private-sector expectations and public policy. In pursuit of this goal, about half of the model's dynamic behavioral equations are explicitly derived as decision rules governing the behavior of representative agents acting with foresight to achieve explicit objectives in the presence of constraints. Forward-looking agents plan in advance to adjust the value of their decision variables to converge over time on target levels, while balancing the cost of being away from the desired level against the cost of adjustment. In particular, the costs of adjusting wages and prices mean that private agents must plan in advance to set out a path for those variables—conditional on expected future demand, relative prices, and policy actions—and the expected future target rate of inflation.

We now demonstrate the basic properties of FRB/US. To do this, we compare the properties of the model and those of the historical data using three methods: inflation and output autocorrelations, impulse responses to a shock to the federal funds rate equation, and the parameters of a dynamic IS-curve equation. To begin, we need to characterize the behavior of monetary policy. To this end, we estimate a parsimonious interest rate reaction function, the results for which are shown in table 1. Like the Taylor (1993) rule, this rule is simple, but it adds some dynamic richness through the inclusion of two terms—one with the change in the output gap, and the other with the lagged federal funds rate.

Given the rule, figure 3 shows the autocorrelations of inflation, measured by the chain-weighted index of personal consumption expenditure prices, and the output gap. In both instances, the asymptotic model moments are represented by the dashed line and the moments estimated from the data are shown as the solid line. The dotted lines are one-standard-error bands for the data-based moments.[19] The model- and data-based moments are computed using the data and equation residuals from the 1980s and 1990s. This

[19]The model moments used here and below are computed from the formula for unconditional covariances. As such, the computed moments correspond to those that would be obtained from a stochastic simulation of infinite length.

Table 1. Estimates of a Simple Monetary Policy Reaction Function (1981:Q1–1999:Q4)

$R = (1 - \alpha_{R1})(\tilde{\pi} + rr^*) + \alpha_{R1} R_{t-1} + \alpha_{y0} y_t + \alpha_{y1} \Delta y_t + \alpha_\pi (\tilde{\pi} - \pi^*)$					
Symbol →	α_{R1}	α_{y0}	α_{y1}	α_π	π^*
Coefficient (Standard Error)	0.84 (0.05)	0.15 (0.05)	0.58 (0.39)	0.25 (0.09)	2.20 (0.60)
$rr^* = 2.5$ (assumed)			SEE = 0.81		

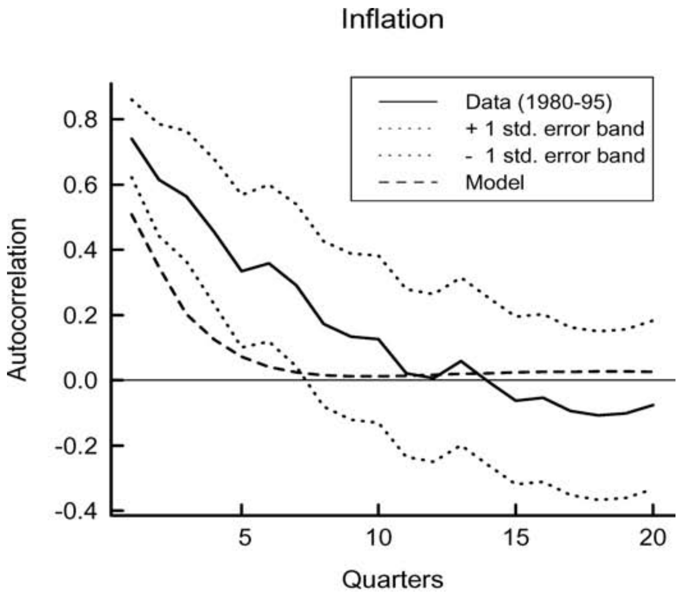
Notes: R is the nominal federal funds rate (quarterly average basis); $\tilde{\pi}$ is the four-quarter rate of change of the (chain-weighted) personal consumption expenditures price index; y is output measured in percent deviations from potential output; rr^* is the steady-state real interest rate (taken to be 2.5 percent); and π^* is the target rate of inflation. Estimation is conducted using instrumental variables with the lagged four-quarter inflation rate, two lags of the output gap, the lagged federal funds rate, and a constant as instruments.

sample period was chosen to approximate a single policy regime and to correspond with the estimated policy rule of table 1.

As emphasized by Fuhrer and Moore (1995), inflation has been very persistent, historically, even during periods in which there were no apparent shifts in the inflation target. Owing to the generalized adjustment cost structure for wages and prices, the FRB/US model also displays a significant amount of inflation persistence, albeit a bit less than is evident in the data. That said, some of the persistence evident in the data is due to the large disinflation that occurred during the first few years of the sample. If the first two years are removed from the sample, the model and data autocovariances are nearly identical at the first four lags. The model also shows considerable persistence in the output gap. Sluggish adjustment of demand components, combined with persistence in the real interest rate driven by the propagation of movements in inflation, drive this result. In this case, autocorrelations from the model and from the data correspond closely.

Now let us consider the impulse response from a disturbance to the estimated federal funds rate. Figure 4 shows the response of the funds rate, inflation, and the output gap. Also shown is the response from a VAR model estimated over the period from

Figure 3. Inflation and Output Persistence (Estimated Rule)



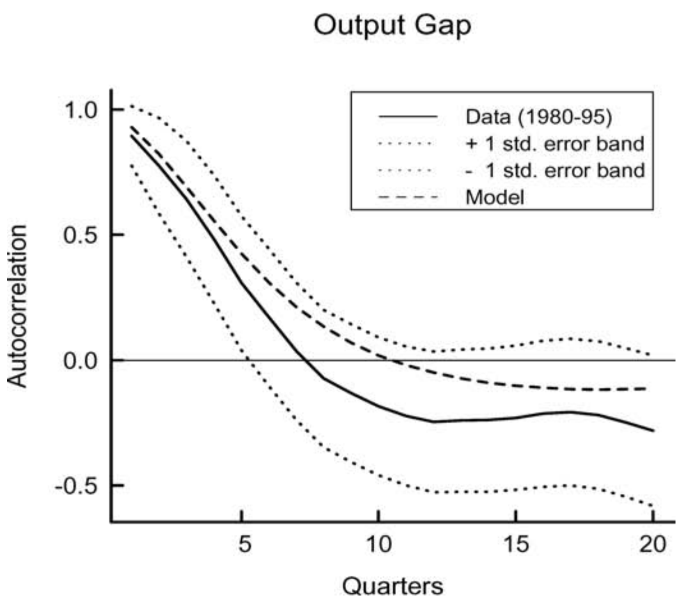
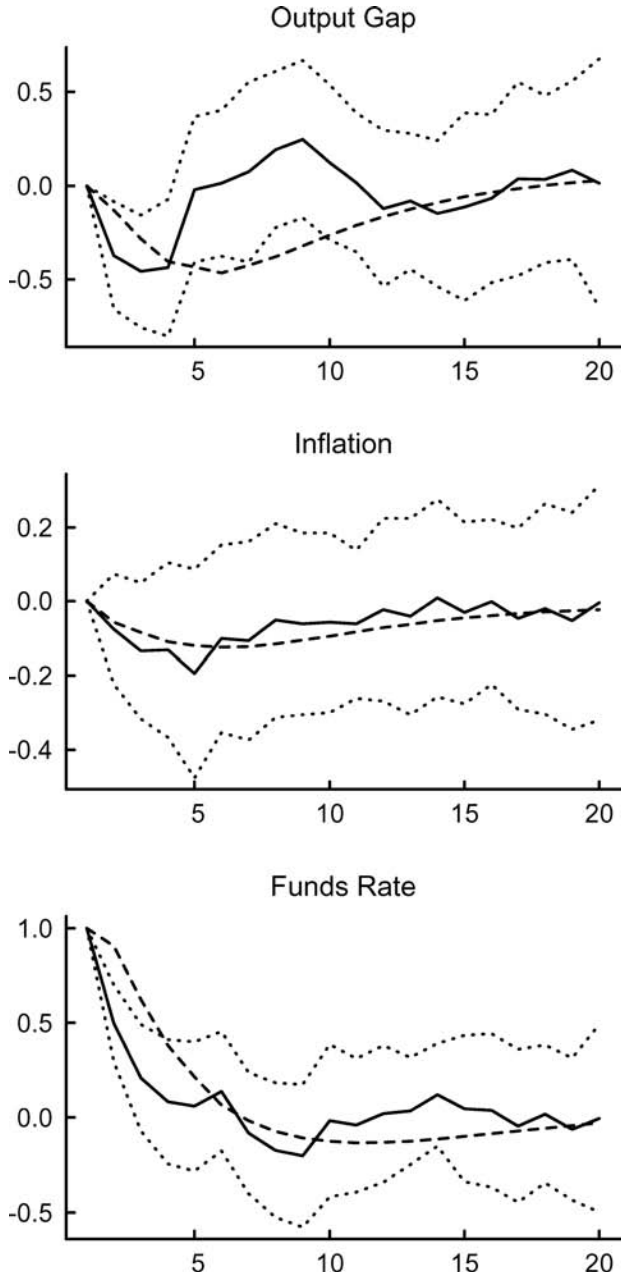


Figure 4. Impulse Response to Funds Rate Shock (100 Basis Points)



1981:Q1 to 1999:Q1, together with one-standard-error bands. The figure shows that the magnitude and the general pattern of output and inflation in FRB/US are close to those of the data.

In addition to the above, we also assessed the model's acceptability by estimating a standard IS-curve equation using model moments, and comparing the model-generated estimates with those from the historical data. The estimates for the model—which can be found in the working paper version of this article—and those in the data were very similar.

In sum, the model fits the data well, both in terms of the individual equations and in terms of its system properties. Furthermore, the model's explicit disentanglement of intrinsic sources of dynamic propagation from expectational sources makes it well suited for the policy analysis experiments to which we shall turn shortly. However, before doing that, we need to discuss the numerical techniques employed to study this issue.

4. Numerical Methods

In all cases studied in this paper, we use a linearized version of FRB/US. Thus, the only source of nonlinearity is the target bands. Setting aside the target bands, the model can be written compactly in companion form and solved employing the algorithm of Anderson and Moore (1985). But while the model itself may be linear, the preceding discussion makes clear that the expectations mechanism for future targets is fundamentally nonlinear. Even in the presence of this nonlinearity, however, there is a way in which we can exploit linear on-line algorithms for solving models.

Given some initial π^*_{t-1}, a vector $E_{t-1}(\pi^*_{t+i}|\pi^*_{t-1}, \sigma, \bar{\pi}^*)$ is computed outside of the model. Using a discrete approximation of the truncated normal distribution, this is possible because $E_{t-1}\pi^*_{t+j}$ is independent of the state variables of the system. Then this nonlinear path is approximated to arbitrary precision with a simple ARMA model, "reverse engineering" the shocks necessary to replicate the nonlinear $E_{t-1}\pi^*_{t+j}$ at each date. The form of the equation stays the same at each iteration of the algorithm; however, the historical shocks to the moving-average errors are changed depending on the initial value of $E_{t-1}(\pi^*_t|\sigma, \bar{\pi}^*)$. An ARMA(1,25) model was chosen, which means that specific values of the shocks can be chosen such that the linear ARMA model represents exactly the nonlinearity up to $j = 25$ and is approximately correct thereafter. In the state-space representation, each additional moving-average term adds an additional equation to the model, so there are trade-offs at work between speed and accuracy.

Some experimentation showed the ARMA(1,25) approximation to be a very good one. The state matrix is augmented with variables for the target and its artificial ARMA determinants by adding a block of dimension 25. The "random shocks" to the target are then selected in such a way that they just so happen to give a nonlinear expected future path for the target. This works because once the form of the recursive system is determined, any sequence of shocks can be supplied without having to bear the computational cost of resolving the model. Since the inherited target rate of inflation is known, the precise sequence of shocks consistent with the correct expected future path for the target can be fed in. The shocks necessary to generate the nonlinear forecast of π^*_{t+i} can be computed ahead of time, indexed against the initial π^*_{t-1} and stored in a grid, which also saves some computing time.

In short, we trick the linear algorithm to give nonlinear forecasts with nonrandom shocks of a particular sequence. By doing this, we can compute what is fundamentally a nonlinear process using linear methods with savings in computational costs that are very large.

5. Results

Our results come from stochastic simulations of the FRB/US model. Each experiment comprises 500 simulations of 200 periods each. After discarding the first 20 observations, we are left with 500 replications times 180 periods, or 90,000 observations per experiment. In most instances, we take the standard deviation of innovations to the target to be 0.25 percent. This is the standard deviation of innovations to the Federal Reserve Bank of Philadelphia's survey of expectations of inflation of ten years ahead; expectations of inflation this far into the future can be reasonably taken to be solely a monetary phenomenon and thus a measure of private agents' expectations of the target rate of inflation. This standard deviation is also close to the standard deviation of changes in the implied target rate of inflation shown in figure 1, for the period since Alan Greenspan assumed the chairmanship of the Federal Reserve in 1987.[20] The width of the bands is ± 1 percentage point, except as noted.

For most of these simulations, monetary policy is governed by equation (2), without error. Thus, at the beginning of each period, agents can infer π^*_{t-1} from last period's federal funds rate setting. Agents then form an expectation of the future path of the target using equations (4) and (5), or (4) and (6), as applicable.

5.1 *Model Properties with Target Bands*

Figure 5 shows the mapping of inflation against the target rate of inflation from our stochastic simulations. (Ignore the dotted lines for the moment.) Two cases are shown. The dashed line is the mapping of inflation against the target when the target follows a pure random walk. This mapping, which we designate $g_{rw}(\pi^*)$, is diagonal: the relationship, on average, between inflation and its target is one-for-one when the target follows a pure random walk.

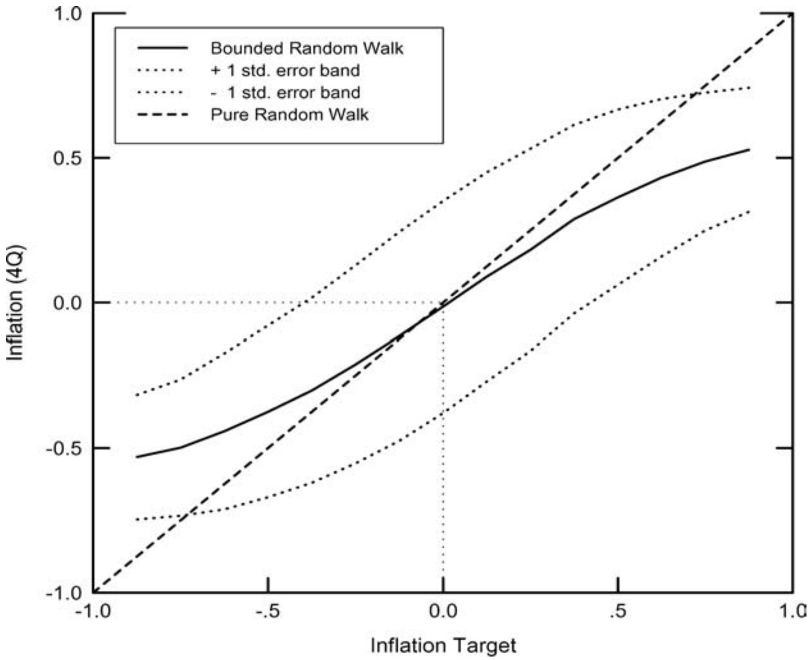
Now consider the solid line. Notice that it is roughly s-shaped. It is, moreover, flatter than g_{rw}, meaning that the range for actual average inflation is less than the range for target inflation. Let us designate this mapping $g(\pi^*) = E_{t-1}\pi(\pi^*|\bar{\pi}^*, \sigma)$, where the right-hand side of the equality notes that this is the expected value of the mapping of inflation on target inflation, conditional on the location of the upper (and lower) bound, and on the size of innovations to the target.[21]

Notice as well that while the bands are between $+1$ and -1 percentage point, the target never actually reaches the edge of the band. For this to happen, one would have to draw exactly the right shock to reach the edge of the band; and the probability of any one particular

[20]It is considerably smaller than innovations in the implicit target rate of inflation for the whole period, even if one omits the discrete shift in the early 1980s from the data. Thus, the magnitude of innovations used here appears to be a reasonable one.

[21]Figure 5 and all subsequent figures are constructed by sorting and ordering observations of π^* from $\underline{\pi}^*$ to $\bar{\pi}^*$, gathering the corresponding observations of π, y, rr, and so forth. The matrix of sorted observations is then grouped into cells of identical widths and averaged across the values of each cell. It is these averages for the cells that are shown in the graphs.

Figure 5. Inflation versus Inflation Target (Shocks to the Target Only)



value is vanishingly small.[22] Comparing this line with the mapping for our unbounded process shows the influence of the credible target band: as π^* gets up close to $\bar{\pi}^*$ (down close to $\underline{\pi}^*$), inflation itself falls below (rises above) the current π^*, on average. The extent to which this occurs is a function of the curvature of expected future target values. In the present case, at the upper end of the curve, where the target rate of inflation is nearing the upper bound, the four-quarter inflation rate averages about 0.55 percent, or more than four-tenths below the target inflation rate. By contrast, when π^* is close to the midpoint, $g(\pi^*)$ is very close to the linear mapping of the pure random-walk case. This outcome illustrates that the establishment of target bands for inflation gives rise to what the exchange-rate

[22] This is a manifestation of our using reflecting barriers. Absorbing barriers would reach the boundary with some probability mass.

target-band literature calls the "honeymoon effect": the promise of fidelity in constraining drift in the target rate of inflation variability is rewarded by restraint in the range of the average inflation rate.[23] As we describe in detail below, this operates through expectations of future target inflation rates impinging on current price- and wage-setting decisions.

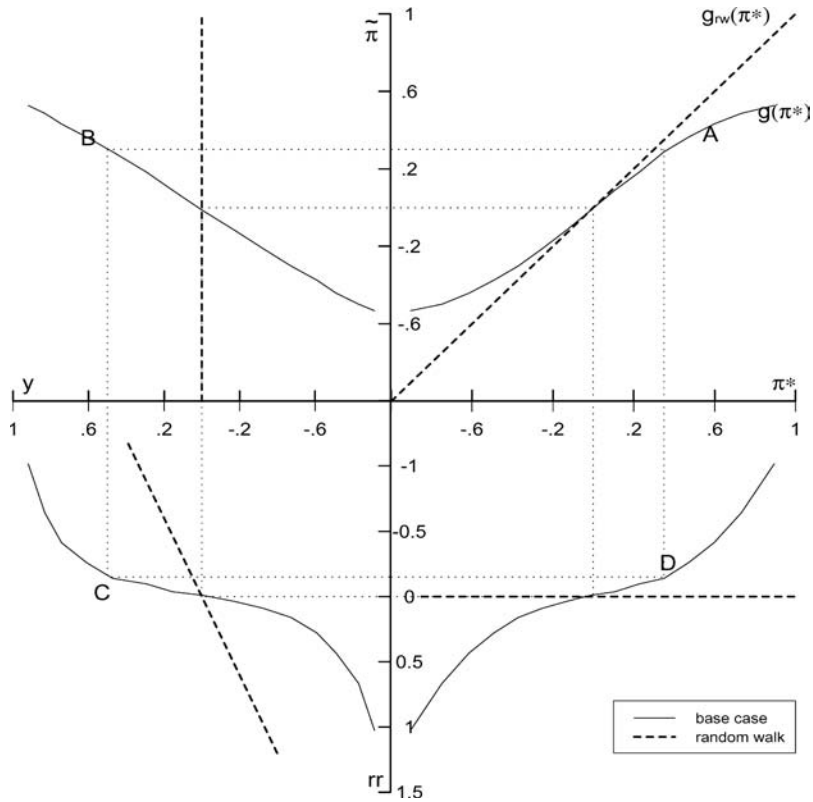
The two dotted lines on figure 5 are the one-standard-error bands for inflation in an economy where there are only shocks to the target rate of inflation. The widths of these bands are not particularly important per se, since the real world is subject to a much wider variety of shocks. What is interesting is the curvature of the confidence bands and their narrowing as the boundaries for the inflation target are approached. What this narrowing indicates is the smaller range of probabilities for future target rates of inflation when the current target is at the edge of the band. Recall that if the inherited rate of inflation were to reach the edge of the band, fully one-half of the distribution of target rates one step ahead would be truncated, reducing the range of future target rates of inflation.

We have established that credible inflation target bands can constrain the range of inflation outcomes, relative to a pure random walk. We now examine some general equilibrium aspects of policy design with credible inflation target bands. To do this, we need to consider a wider range of variables, which we can do with the aid of figure 6.

Before explaining the economics behind the lines in figure 6, let us first just describe what the four quadrants represent. The upper-right quadrant of this four-quadrant diagram shows $g(\pi^*|\bar{\pi}^* = 1)$ and so merely repeats figure 5—except without the error bands. Moving to the upper-left quadrant, we have the mapping of inflation on the output gap. This is *not* the slope of a Phillips curve as such; rather, it is the (mean of the) stochastic relationship between the target inflation rate and the actual inflation rate. The lower-left quadrant shows the mapping of output and the real federal funds rate, both measured in deviations from steady-state levels. Finally, the lower-right quadrant completes our macroeconomic tour with the mapping of the real federal funds rate against the target rate

[23] See, e.g., Krugman (1991).

Figure 6. Summary of Macroeconomic Outcomes (Random Walk versus Inflation Target Bands)



of inflation. This four-quadrant graph portrayal allows us to trace the average levels of four variables at once, as illustrated by the two cases marked by the dotted lines.

Now let us walk through the economics of figure 6. Our base case is shown as the thin solid line, with the case of the pure random walk shown by the dashed line. We begin with point A on the base-case mapping $g(\pi^*)$ in the upper-right quadrant. At this point, $E_{t-1}(\pi_t|\pi_t^*) < \pi_t^*$ because private agents know that the target is more likely to fall than rise in the future. Because firms find it costly to adjust prices and wages, firms must plan ahead to minimize costs, and so this expectation influences their price- and wage-setting behavior *today*, and thereby affects realized inflation today.

Moving along the dotted line to point B in the upper-left quadrant, we see that point A corresponds with an output gap that is positive, on average. We see that the relationship between the average level of output and inflation is positively sloped and linear in the base case, while in the random-walk case there is no output gain to be had from allowing drift in the target. The random-walk case is simply an outcome of the superneutrality of the model: there is no trade-off between output and inflation and thus no slope of the dashed line in the upper-left quadrant. For the base-case solution, the "bias" in inflation expectations relative to the current target rate of inflation means that output can be positive because inflation is expected to fall in the future.[24] To see why this is so, we need to move to the lower-left quadrant.

In the lower-left quadrant, we observe a negative relationship between the output gap and the average level of the real interest rate. Notice that at point C, the real federal funds rate is slightly *negative* even though output is substantially positive. This is feasible because the "bias" in expected inflation—which carries a large weight of 3 in the policy rule—offsets the positive output gap. To put the same idea a different way, the lower expected future inflation rate at point A means that there can be a lower real interest rate at the corresponding point C, which supports higher output at point B.

Note that the curved line for the base-case model implies that the quasi-steady-state interest elasticity changes nonlinearly with the target rate of inflation. The shape of the curve comes from

[24] It is also useful to think of this from the monetary authority's point of view. The authority is attempting to move inflation to the current-period target shown at point A. But because the expected future target is lower than the current target, current inflation is lower than the current target, on average. By contrast, in the random-walk case, inflation, on average, will be at the target value. This means that policy will be easier at target rates of inflation like point A with bands than it would be when the target follows a random walk. This we see in point C in the lower-left quadrant. The easier policy, in turn, implies an average positive output gap, as shown by point B in the upper-left quadrant.

the cumulative normal distribution. In the random-walk case, it is constant.[25]

Finally, moving to point D in the lower-right quadrant, we are reminded that the real interest rate is generally negatively related to the target rate of inflation. In the pure random-walk case, the horizontal dashed line shows that there is no relationship, on average, between the target rate of inflation and the real interest rate. At point D, the real interest rate is lower than it would be if agents did not believe the target was bounded. The expectation of a lower target rate of inflation in the future allows the current real interest rate to be lower than it otherwise would be, thereby supporting higher output than usual.

In sum, we see that the implications for the monetary authority of credible target bands for inflation are (i) a narrower range of inflation for a given range of the inflation target; (ii) a positive range of output, meaning that there exists a trade-off between output and inflation (although it cannot be said that this is a long-run trade-off); and (iii) a range for real interest rates.

5.2 Quantitative Performance

The choice of target band limits of ± 1 percentage point was an arbitrary one. To explore the implications of alternative assumptions, table 2 shows the effect on aggregate (unconditional) volatility of alternative choices. The statistics are based on stochastic simulations using the full set of stochastic shocks and pooling over all simulated observations. The first column in the body of the table corresponds to a credible fixed target rate of inflation, the usual assumption undertaken in assessments of policy rules. The last column covers the pure random-walk case. In that case, the values of many simulated variables are unbounded and are shown as "n/a."[26]

[25] We show the interest elasticity of aggregate demand in the random-walk case to be the -0.33 value computed in table 1.

[26] In this case the range is for consumer price inflation as measured by the four-quarter growth rate in the chain-weighted PCE price index. All the countries that have announced specific targets for inflation have named a twelve-month rate of some consumer or retail price index as the target variable. The FRB/US model does not have a CPI and is a quarterly model. Thus, we use the growth rate of the four-quarter PCE price index here.

Table 2. Simulated Standard Deviations of Target-Band Regimes

Column No. → Endogenous Variable	Bound on Inflation Target (1) $\pi^* = 0$	(2) $\pi^* \in [\pm 0.5]$	(3) $\pi^* \in [\pm 1.0]$	(4) $\pi^* \in [\pm 1.5]$	(5) $\pi^* \in [\pm \infty]$
R	2.76	2.87	3.01	3.11	n/a
y	2.70	2.74	2.90	2.98	3.10
$\tilde{\pi}$	0.88	0.93	1.04	1.21	n/a

Notes: See the notes to table 1 for variable mnemonics. The column heading $z \in [\pm x]$ indicates that z is bounded within values of $-x$ and $+x$, inclusively. Numbers in the rows of the body of the table are standard deviations computed from pooling 500 draws of 180 periods each, after discarding twenty observations per draw to account for initial conditions. In all cases, fifty-one stochastic shocks are drawn using the variance-covariance matrix of the period from 1981:Q1 to 1995:Q4.

The most remarkable observation from table 2 is how unremarkable the differences are: tighter target-band limits improve performance of the economy measured in terms of the unconditional standard deviation of output, inflation, and the funds rate, but not by a great deal. Furthermore, the gains get smaller as one moves from right to left in the table.

At the outset of this paper, we noted that among the devices through which inflation-targeting countries attempt to commit to constraining their discretion are *ranges for inflation itself* (as opposed to bands for target inflation).[27] In table 3 we examine some aspects of the efficacy of ranges for inflation as a device for containing discretion, or policy drift, by looking at the likelihood of inflation rate departing from selected ranges. These ranges are shown in the far left column of the table. Observe that they refer to *ranges* for

[27] It is interesting in this regard to note the motivation for ranges for inflation as given by John Crow, then Governor of the Bank of Canada (1991, 11): "The purpose of setting out formal targets is to provide a clear indication of the downward path for inflation over the medium term so that firms and individuals can take this into account in their economic decision-making. . . . [If people] base their economic decisions on this declining path for inflation, the objectives can be readily achieved and will contribute to lower interest rates. The inflation targets also provide information . . . [that] should provide a better basis than before for judging the performance of monetary policy."

Table 3. Likelihood and Duration of Inflation Target Range Violations

Column No. → **Inflation Range**		**Bound on Inflation Target** (1) $\pi^* = [0]$	(2) $\pi^* \in [\pm 0.5]$	(3) $\pi^* \in [\pm 1.0]$	(4) $\pi^* \in [\pm 1.5]$	(5) $\pi^* \in [\pm \infty]$
$\lvert\tilde{\pi}\rvert > 2.0$	Likelihood	0.03	0.03	0.05	0.10	0.45
	Duration	2.10	2.20	2.50	3.00	n/a
$\lvert\tilde{\pi}\rvert > 1.5$	Likelihood	0.10	0.11	0.15	0.22	0.57
	Duration	2.60	2.60	3.00	3.80	n/a
$\lvert\tilde{\pi}\rvert > 1.0$	Likelihood	0.28	0.28	0.34	0.43	0.70
	Duration	3.20	3.30	3.90	4.90	n/a

Notes: "Likelihood" is the average proportion of quarters in which $\tilde{\pi}$ exceeded the absolute value of the range for inflation, measured relative to the long-run target. "Duration" is the average length of time $\tilde{\pi}$ spends outside of the range, once a violation is recorded. Computations come from averages across the 500 draws for 180 dates in each draw. See also the notes to table 2.

inflation itself. We examine three ranges for inflation—$\tilde{\pi}$, without the asterisk but with a tilda overstrike to indicate that it is the four-quarter rate. To keep the syntax clear, we shall refer to desired variability of inflation itself as *ranges*, and limits on the inflation target as *bands*. An examination of this sort gives an idea of the extent to which commitment to a range of inflation (an observable variable, but arguably not controllable) can stand in for commitment to a band width on target inflation (an unobservable variable, but controllable).

Across the top of the table, we show widths for the band for target inflation. In the far left column of the table are shown three ranges for inflation variability. The body of the table shows, therefore, the bands on inflation targets that would have to be kept to in order to deliver the performance on four-quarter inflation ranges that is shown. Performance in this regard is shown in two dimensions—the likelihood of being outside the inflation range shown at left measured in terms of the proportion of quarters outside the range, and the duration of the average range violation in quarters. As a concrete example, the first row of column 1 of the table shows that with no variability of the inflation target at all, inflation will depart from a range for inflation of ± 2 percentage points about 3 percent of time. The 2.10 number immediately below says that these departures tend to last only about six months. The last row of the same column shows that even with a fixed target rate of inflation—and central bank preferences that weight inflation control fairly highly—the authority should not expect to keep inflation within ± 1 percentage point of the target much more than about 70 percent of the time.

Reading the table from right to left across any given row shows that as the range of target inflation rates is narrowed, the likelihood and duration of inflation ranges diminishes. This much is hardly surprising. A more interesting observation is that there are sharply diminishing returns, regardless of the width of the inflation range. For example, for an announced inflation range of ± 1.5 percentage points (the third row), cutting the target band width from ± 1 percentage point (column 3) to .5 percentage point (column 2) reduces the frequency of violations from 0.15 to 0.11. A further cut of the target band width to zero (column 1) reduces the frequency only a trivial amount.

In general, there is a notably high frequency of range violations for all target band widths and for most inflation ranges: to get the likelihood of range violations below 10 percent one must accept very wide inflation ranges of ±2 percentage points. Economically speaking, ranges this wide have little meaning for an economy like the United States. Focusing on ranges for inflation of ±1 percentage point—that is, on the range that most inflation-targeting countries use—we see a substantial likelihood of violations, and departures usually last more than three quarters when they occur.

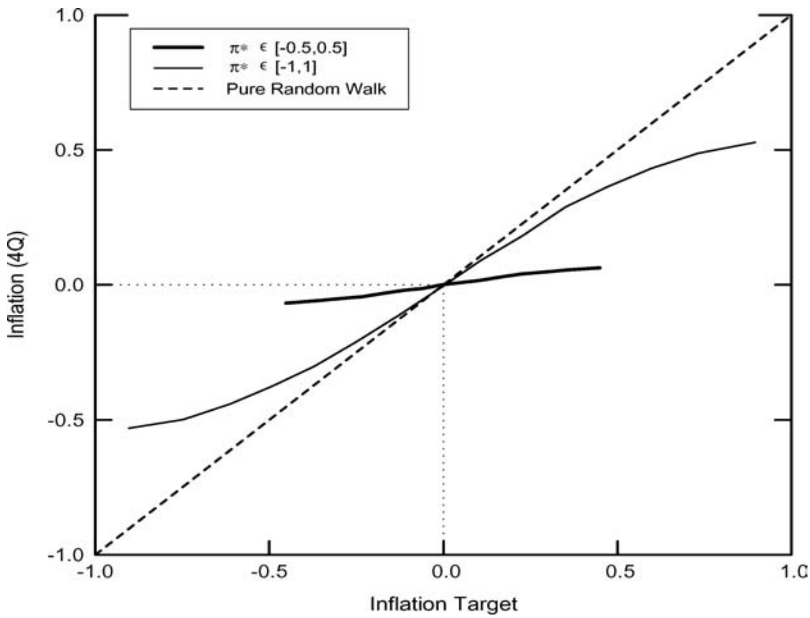
More than anything else, this reflects the fact that a substantial portion of the variability of inflation comes from shocks to the wage-and-price block of the model. There is little that the monetary authority can do to offset such shocks. Blackburn and Christensen (1989) emphasize that monetary policy credibility comes from a combination of the *willingness* of the monetary authority to pursue its inflation objective and the *ability* of the authority to achieve the objective. Table 3 suggests that there are limits to what the Federal Reserve can expect to deliver in terms of outcomes of monetary policy, even if bounds on inflation targets are specified, kept, and believed. On the positive side, however, the table also shows that reductions in the width of the inflation target bands below ±.5 percentage point yield very little in terms of the frequency of range violations, regardless of the range selected.[28]

5.3 Comparative Statics

The unconditional standard deviations shown in table 2 mask some considerable variation in the conditional statistics. Figure 7 replicates figure 5 in showing the mapping from the inflation target to inflation, this time for target-band limits of ±.5 percentage point

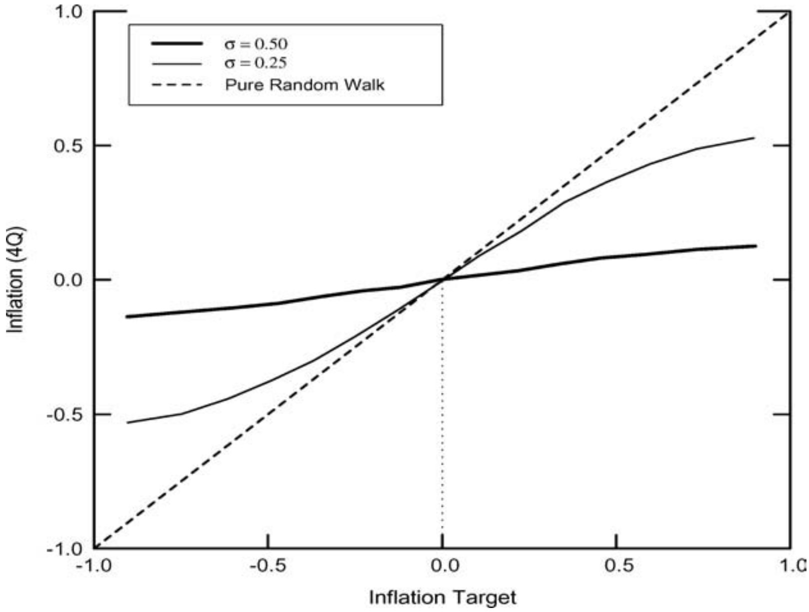
[28]There is also the prospect, alluded to by Chairman Greenspan in the quote at the beginning of this paper, that some of the "exogenous shocks" to the model's wage-price blocks might be expectations scares brought about by low credibility, on average, through history. This would be the case, for example, if policy control were so languid as to induce sunspot equilibria; see, e.g., Clarida, Galí, and Gertler (1998). To the extent that this is true, the variance of these "supply shocks" will be lower in the future than in the past, the honeymoon effect of figure 5 will be larger, and the probability of inflation-range violations will be lower than what is shown in table 3.

Figure 7. Inflation versus Inflation Target (Selected Band Widths)



(the thick line) together with our base-case assumption of ± 1 percentage point (the thin line). The differences here are much more pronounced than table 3 might have led one to believe; the honeymoon effect is apparent and substantial at all ranges of $\pi^* \neq 0$, not just beyond the first couple of tenths of a percentage point away from the midpoint of the range. Even in the presence of the unconditional inflation variability of 0.93 (table 2, third row, column 2), this is an economically important difference in performance. By the same token, however, a widening of the target band produces a flattening in $g(\pi^*)$ and a region near the midpoint of the bands in which $g(\pi^*)$ and g_{rw} are observationally equivalent. It follows that fidelity in controlling inflation target drift would not necessarily be rewarded with rising credibility and hence a honeymoon effect. This observation goes some way in explaining why countries that target inflation expend so much time and effort communicating their policy intentions. They hope to reverse the causal order of good targeting

Figure 8. Inflation versus Inflation Target (Selected Standard Deviations of Target Shocks)



performance leading to enhanced reputation and thus higher credibility with public declarations so that the promised improvement is easier to achieve in the first place.[29]

The other parameter that affects the honeymoon effect is the magnitude of innovations to the target rate of inflation. In our base case, the assumed standard deviation of disturbances to π^* is 0.25 percentage point. Figure 8 shows the implications of higher variability in the monetary authority's target, holding constant the

[29]That inflation targeting with bands is an attempt to get a "free lunch" along the lines described in the text is made clear for the case of Canada by Freedman (1995, 27): "These ranges were in fact smaller than were called for in empirical work done at the Bank. There is a trade-off . . . [t]he wider are the bands, the higher is the probability of successful achievement of the targets but the less useful are the targets in changing behavior. In the end, we decided to use somewhat narrower bands to avoid the problem that overly wide bands might leave the impression that the authorities were not serious about bringing inflation down." In fact, violations of the bands have been much less common than the Bank's prior empirical would have suggested was possible.

target band width at ±1 percentage point. As discussed above, a larger σ implies a more substantial nonlinearity in expected future target values for any initial $\pi^* \neq 0$. Not surprisingly, therefore, figure 8 shows $g(\pi^*|\sigma = 0.5)$ to be flatter than in the base case. With the larger σ, there is a honeymoon effect even with relatively wide target bands. This shows that a monetary authority with erratic preferences has considerably more to gain from the establishment of target band widths, if such an authority can secure credibility. It also may help explain why those countries that have adopted inflation targets with bands have done so after abandoning under duress an exchange-rate targeting regime, while those countries that have announced imprecise inflation-targeting objectives do not have a record of such difficulties.[30]

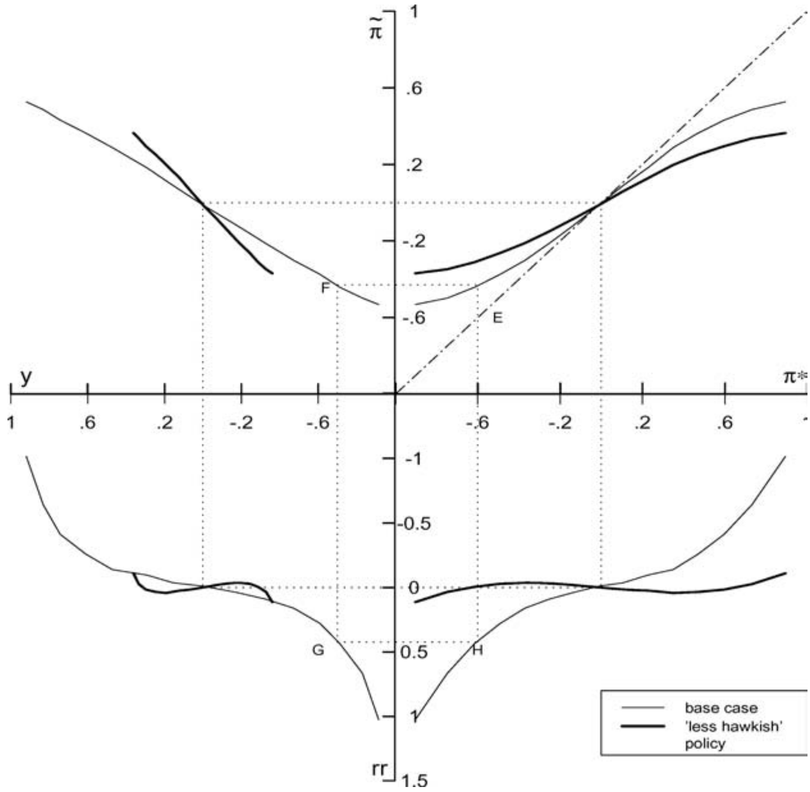
5.4 Alternative Policies

Let us now compare our base-case results with those of a few alternatives, beginning with a less hawkish policy. Recall that for our base case, we assumed that the monetary authority placed three times the importance on inflation stabilization as on output stabilization. We now consider a policy rule derived with the opposite loss function weights: 0.75 on output and 0.25 on inflation. This raises the coefficient on the output gap of 0.9 in our base-case rule to about 1.1 and lowers the coefficient of the twelve-quarter moving average of inflation from 3 to approximately 0.75.

The thin lines in figure 9 are the same lines as in figure 6; they represent our base case. The thick lines are computed in exactly the same way as the thin lines, except using the coefficients of the less hawkish rule. Starting, once again, in the upper-right quadrant, we see that $g_{lh}(\pi^*)$ is flatter than it is in the base case. This means that there is a larger honeymoon effect, or to put the same point a different way, there is less variability in inflation for a given degree of variability in the target than in the base case. In the upper-left quadrant, we see a steeper relationship between the average level

[30] Canada is the only country whose decision to adopt a well-defined inflation-targeting regime was not prompted by a crisis. In most cases, inflation-targeting countries had previously failed to maintain an exchange-rate targeting regime.

Figure 9. Implication of Different Policy Preferences Under Inflation Target Bands



of output and inflation, and, more importantly, a much narrower range of output gaps is implied by that relationship. In the lower-left quadrant, we see that the range of real interest rates associated with these output gaps is not only smaller, but a slight backward bend appears: a decline in the output gap, beginning from zero, arising from a decline in the target that is accompanied by a smaller decline in inflation, results in a (very small) decline in the real interest rate.

The quantitative assessment of this rule is shown in table 4. Each column in the body of the table records aspects of the performance of a particular policy rule: the first column shows the performance of our base-case rule; column 2 shows the performance of the less hawkish rule. (We will ignore column 3 for now.) The upper panel

Table 4. Simulated Standard Deviations of Target-Band Regimes for Alternative Policies

Column Number → **Policy Rule in Use →**	**(1)** **Base Case**	**(2)** **Less Hawkish**	**(3)** **Forward Looking**
y Standard	2.90	2.04	2.75
$\tilde{\pi} - \pi^*$ Deviations	0.99	1.40	1.05
$R - \pi^*$	3.07	2.78	3.79
Preferences	**Normalized Losses**		
Base Case: $\varphi = 0.75$	1.00	0.88	0.96
Less Hawkish: $\varphi = 0.25$	1.81	1.00	1.65
	Range Violations $(\pm\bar{\pi} = 1)$		
Likelihood	0.15	0.29	0.16
Duration	3.00	4.90	3.20
Loss Function: $\sum_{i=1}^{180} \varphi[\tilde{\pi}_{t+i} - \pi^*_{t+i}]^2 + (1-\varphi)y^2_{t+i}$			

of the table shows the standard deviations of this rule for the base-case parameterization of the bounded random walk. As one would expect, the standard deviation of output is lower for the less hawkish rule than for the base-case policy rule, while the standard deviation of inflation is higher.

The middle panel of the table is more interesting. There we record the loss associated with each rule indexed against the preferences that generate the rules. In each case, we have normalized the losses for the two rules so that, for example, the base-case rule in column 1 is assessed a (normalized) loss of unity for base-case preferences. Similarly, in column 2, the less hawkish rule carries a normalized loss of unity for less hawkish preferences. The interesting result is in the other two entries in these two rows of this panel: notice in column 1 that the performance of the base-case rule under less hawkish preferences is poor: there would be an 81 percent deterioration in performance from the perspective of a less hawkish monetary authority from adopting the base-case rule as an operating procedure. Now examine the performance of the less hawkish rule under base-case preferences in the first column. With a value shown there of 0.88, we see that *the base-case policymaker actually prefers the less hawkish rule.*

The result shown is not an error. It arises because the so-called optimal rule for base-case preferences is actually only optimal *conditional on a fixed target rate of inflation* (or a random walk).[31] In the presence of a bound on the random walk, the rule is no longer optimal. This is because the credibility of the target bands does much of the work for the authority; given the bands, adding tough inflation targeting to the policy mix only makes matters worse. To see why this arises, consider point E on the thin line in the upper-right quadrant of figure 9. Observe that the level of output that corresponds with this, point F in the upper-left quadrant, is low. With output low, all else equal, the real rate should be low as well. The reason why the real rate is *not* low is that the monetary authority is trying to move $\tilde{\pi}$ *down* to π^*, the current target rate of inflation. However, private agents know that π^* is likely to be higher in the future. The inflation rate that arises reflects this expectation on the part of price and wage setters who, after all, must consider current and future costs of adjusting prices: it is optimal for them to avoid reducing prices today if they would only have to raise them again in the near future.

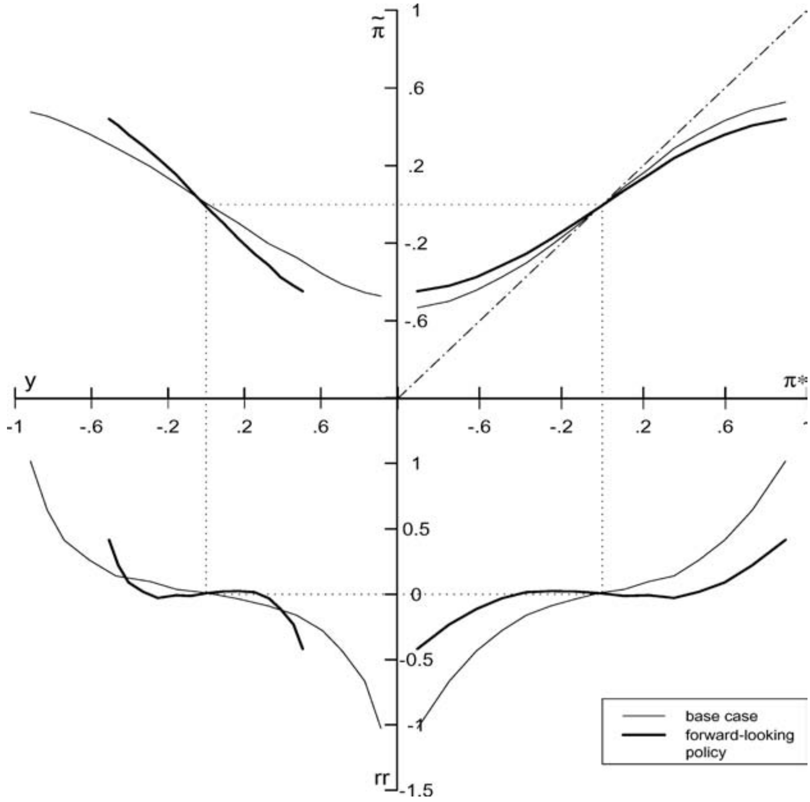
But the monetary authority is working against the very effect it has created by adopting target bands in the first place. It follows that a rule that puts less weight on inflation control, in recognition that private agents are doing some of the controlling of prices through their expectations, will do better on average.

There are other policy recipes that outperform our base-case rule. Figure 10 provides one example. The thin line in figure 10 is, once again, our base case. The thick line represents a rule that is forward looking in inflation. In particular, we have replaced the twelve-quarter moving average of current and past inflation less the current target in equation (2) with a four-quarter moving average of current and future inflation less the *expected* future inflation target:

$$R_t = rr^* + \sum_{i=0}^{3} \pi_{t-i}/4 + 3.0 \left[\sum_{i=0}^{3} \left(\pi_{t+i} - \pi^*_{t+i} \right) \right] + 0.9 y_t. \quad (7)$$

[31]The base-case rule is optimal for the case of a pure random walk, provided that the loss is computed around the (possibly drifting) target rate of inflation, since a fixed target and a pure random walk are identical in expected value terms.

Figure 10. Implication of Forward Looking Policy Under Inflation Target Bands



In our base-case scenario, the monetary authority is conducting policy myopically, carrying out short-term target-seeking actions without regard to how the longer-term process of determining the future of the target will play itself out. With the forward-looking rule, the central bank takes into consideration expected future target changes, even if it is not solving an optimization problem for the future target simultaneously with the determination of the coefficients of the rule.

The results in this case are qualitatively similar to those in figure 6. A forward-looking monetary authority recognizes that private agents know that the target rate of inflation will regress toward the midpoint of the bands over time. By acknowledging this fact in making policy decisions, the monetary authority can improve average policy outcomes in terms of both output and inflation. The middle panel of table 4 shows us that based on the loss calculation shown, the base-case monetary authority would once again prefer the forward-looking rule to the base-case rule, if only by a slight amount. It would be easy to make too much of this comparison, however, since the real federal funds rate variability is significantly higher for the forward-looking rule than for the base-case rule.

We leave this section of the paper, and this topic, with two observations. First, under the assumptions of this section, it is possible to improve on the economic performance of the simple monetary policy rules that depend, linearly and axiomatically, only on a small set of past values. If monetary policymakers' targets do shift over time but are bounded in some way, and if agents believe in these bounds, then private agents' expectations will be biased, ex post, in a way that is completely consistent with the economic environment. An optimizing policy authority that chooses to ignore such biases will forgo improvements in average policy outcomes that can be accrued by encompassing a wide range of indicators of expected future inflation. Similarly, since expectations of target movements on the part of private agents are doing part of the work of the authority in directing price-setting behavior toward long-term monetary policy goals, a Pareto-improving response of the authority is to substitute away from inflation stabilization and toward output stabilization. Second, our analysis offers a rationale for the passivity in monetary policy operations that many observers have noted appears to be a characteristic of policy in the United States.[32] To see this, examine the thick lines associated with the forward-looking policy rule in the two right-hand-side quadrants of figure 10. The honeymoon effect implies that both inflation and the real interest rate are much flatter than they would be under the random-walk case. Because expectations are being anchored by the credible target band, there is less impetus to adjust the nominal interest rate as the target inflation rate moves over time. Essentially, the public is doing part of the Federal Reserve's work for it. We hasten to add, however, that it would be easy to take this argument too far: the lines in our figures represent

[32]See, in particular, Huizinga and Eijffinger (1999), Rudebusch (1999), and Woodford (1999).

averages across hundreds of business cycles. Within those cycles, policy is not "passive" at all. Still, this observation, combined with the results portrayed in figure 9 for the less hawkish rule, does provide some insights on these issues.

6. Concluding Remarks

This paper has considered the implications of uncertain policy targets for the conduct of monetary policy. In particular, we have modeled target variability as either a random walk or a bounded random walk, and examined the implications of the bounds for policy and economic outcomes. Uncertain policy targets were argued to be a manifestation of either discretionary policy or the unwillingness of monetary authorities to respond to disturbances in a timely fashion.

Conditional on a given policy rule, we found that credible constraints on the variability of the target rate of inflation can reap benefits in reduced inflation variability, without increasing output variability, in the same channels as Chairman Greenspan outlined at the outset of this article. However, the improvements are diminishing as band width gets narrower and narrower. All else equal, this provides modest support for point targets for inflation as opposed to Australia-style ranges.

More important, consistent with popular wisdom, we found that a central bank that enjoys credibility in the range of target drift can exploit that credibility by substituting, at the margin, away from inflation control and toward output stabilization. This arises because the expectation that the central bank is likely to contain drift in the target creates a nonlinearity in agents' expectations. This pins down expectations of future inflation on the part of firms and workers, and thereby constrains their inclination to respond to shocks by touching off a wage-price spiral. It is this pinning down of future inflation that allows the monetary authority to be less aggressive in its inflation-control policies than would be the case with either a random-walk target or a fixed target. Taken together, the results here make a case for favoring time-invariant point targets for inflation, as opposed to target ranges, or comfort zones.[33]

[33] See Mishkin (2008) for a spirited and multifaceted argument in favor of point objectives, in part along the lines of this paper.

This finding is reminiscent of Orphanides and Williams (2005), who find that if the central banks can alleviate the need of private agents to learn the target rate of inflation, the resulting anchoring of inflation allows an easier policy regime than would be the case.[34]

The question remains as to how a central bank might establish credible constraints on inflation targets. We noted in our discussion of figure 7 that for some band widths it is hard to demonstrate policy fidelity and hence hard to build reputation through performance. One way to overcome this is to employ public announcements of a fixed inflation target and a range for inflation outcomes like most inflation-targeting countries have done. Table 4 of this paper described how doing so can direct agents' expectations toward the ultimate objective of policy, while still permitting some time variation in policy setting. A central bank that starts out with a reputation for allowing its inflation target to drift can benefit from a "honeymoon effect" wherein the bank reaps the reward of decreases in average inflation variability that are greater than the restriction in the variability of the target.

References

Aksoy, Y., A. Orphanides, D. H. Small, V. Wieland, and D. A. Wilcox. 2006. "A Quantitative Exploration of the Opportunistic Approach to Disinflation." *Journal of Monetary Economics* 53 (8): 1877–93.

Anderson, G., and G. Moore. 1985. "A Linear Algebraic Procedure for Solving Linear Perfect Foresight Models." *Economics Letters* 17 (3): 247–52.

Bank of Canada. 1991. "Background Note on the Targets." Released on February 26, 1991. Reprinted in *Bank of Canada Review* (March 1991).

Bernanke, B. S. 2005. "The Economic Outlook." Remarks at a Finance Committee luncheon of the Executives' Club of Chicago, Chicago, IL, March 8.

[34] As a referee has pointed out, a useful extension of the current paper would be to drop the assumption that agents know the time-varying inflation target with a one-period lag and instead have agents formally learn the target over time. Our conjecture is that this would magnify the findings herein.

Blackburn, K., and M. Christensen. 1989. "Monetary Policy and Policy Credibility: Theories and Evidence." *Journal of Economic Literature* 27 (1): 1–45.

Blinder, A. S. 1998. *Central Banking in Theory and Practice.* Cambridge, MA: MIT Press.

Bomfim, A. N., 1997. "The Equilibrium Fed Funds Rate and the Indicator Properties of Term-Structure Spreads." *Economic Inquiry* 35 (4): 830–46.

Bomfim, A., R. Tetlow, P. von zur Muehlen, and J. Williams. 1997. "Expectations, Learning and the Costs of Disinflation: Experiments Using the FRB/US Model." Finance and Economics Discussion Series, No. 1997-42, Board of Governors of the Federal Reserve System.

Brayton, F., E. Mauskopf, D. Reifschneider, P. Tinsley, and J. Williams. 1997. "The Role of Expectations in the FRB/US Macroeconomic Model." *Federal Reserve Bulletin* 84 (4, April: 227–45.

Brayton, F., and P. Tinsley (eds.) 1996. "A Guide to FRB/US: A Macroeconometric Model of the United States." Finance and Economics Discussion Series, No. 1996-42, Board of Governors of the Federal Reserve System.

Clarida, R., J. Galí, and M. Gertler. 1998. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory." NBER Working Paper No. 6442.

Dixit, A. 1993. *The Art of Smooth Pasting.* Fundamentals of Pure and Applied Economics, Vol. 55. Langhorne, PA: Harwood Academic.

Freedman, C. 1995. "The Canadian Experience with Targets for Reducing and Controlling Inflation." In *Inflation Targets*, ed. L. Leiderman and L. E. O. Svensson, 19–31. London: Centre for Economic Policy Research.

Fuhrer, J., and G. Moore. 1995. "Inflation Persistence." *Quarterly Journal of Economics* 110 (1): 127–59.

Greenspan, A. 1998. *The Federal Reserve's Semiannual Monetary Policy Report.* Testimony of Chairman Alan Greenspan before the Subcommittee on Domestic and International Monetary Policy of the Committee on Banking and Financial Services, U.S. House of Representatives, February 24. Washington, DC: Board of Governors of the Federal Reserve System.

Huizinga, H., and S. Eijffinger. 1999. "Should Monetary Policy Be Adjusted Frequently?" CEPR Discussion Paper No. 2074.

Johnson, N. L., S. Kotz, and N. Balakrishnan. 1994. *Continuous Univariate Distributions*, Vol. 1, 2nd ed. New York: Wiley.

Kohn, D. L. 1991. "Commentary: The Federal Reserve Policy Process." In *Monetary Policy on the 75th Anniversary of the Federal Reserve System*, ed. Joseph Belongia, 96–103. Boston: Kluwer Academic.

Krugman, P. R. 1991. "Target Zones and Exchange Rate Dynamics." *Quarterly Journal of Economics* 106 (3): 669–82.

Meltzer, A. H. 1991. "The Fed at Seventy-Five." In *Monetary Policy on the 75th Anniversary of the Federal Reserve System*, ed. Joseph Belongia, 3–65. Boston: Kluwer Academic.

Meyer, L. H. 1997. "Monetary Policy Objectives and Strategy." *Business Economics* 32 (1): 17–21.

———. 1998. "Come with Me to the FOMC." Remaks at the Gillis Lecture, Willamette University, Salem, OR, April 2.

Mishkin, F. S. 2008. "Comfort Zones, Shmumfort Zones." Speech at the Sandbridge Lecture of the Virginia Association of Economists and the H. Parker Willis Lecture of Washington and Lee University, Lexington, VA, March 27.

Orphanides, A., and V. Wieland. 2000. "Inflation Zone Targeting." *European Economic Review* 44 (7): 1351–87.

Orphanides, A., and D. W. Wilcox. 2002. "The Opportunistic Approach to Disinflation." *International Finance* 5 (1): 47–71.

Orphanides, A., and J. C. Williams. 2005. "Imperfect Knowledge, Inflation Expectations, and Monetary Policy." In *The Inflation-Targeting Debate*, ed. B. S. Bernanke and M. Woodford, 201–34. NBER Studies in Business Cycles, Vol. 32. Chicago: University of Chicago Press.

Reifschneider, D., R. Tetlow, and J. C. Williams. 1999. "Aggregate Disturbances, Monetary Policy, and the Macroeconomy: The FRB/US Perspective." *Federal Reserve Bulletin* 85 (1): 1–19.

Riboni, A., and F. J. Ruge-Murcia. 2008. "The Dynamic (In)efficiency of Monetary Policy by Committee." *Journal of Money, Credit, and Banking* 40 (5): 1001–32.

Rudebusch, G. D. 1999. "Is the Fed Too Timid? Monetary Policy in an Uncertain World." Unpublished Manuscript, Federal Reserve Bank of San Francisco.

Siklos, P. L. 1997. "Charting a Future for the Bank of Canada: Inflation Targets and the Balance between Autonomy and Accountability." In *Where Do We Go From Here: Inflation Targets in Canada's Monetary Policy Regime*, ed. D. E. W Laidler. Policy Study No. 29. Toronto: C. D. Howe Institute.

Taylor, J. B. 1993. "Discretion versus Policy Rules in Practice." *Carnegie-Rochester Conference Series on Public Policy* 39 (1): 195–214.

Williams, J. C. 2003. "Simple Rules for Monetary Policy." *Economic Review* (Federal Reserve Bank of San Francisco) 2003: 112.

Woodford, M. 1999. "Optimal Monetary Policy Inertia." NBER Working Paper No. 7261.

Yellen, J. L. 2006. "Enhancing Fed Credibility." Speech at the Annual Washington Policy Conference, National Association for Business Economics, Washington, DC, March 13.

Central Bank Policy Rate Guidance and Financial Market Functioning*

Richhild Moessner[a] and William R. Nelson[b]
[a]Bank for International Settlements
[b]Federal Reserve Board

Several central bankers have expressed concern that providing forecasts of future policy rates may impair financial-market functioning. We look for evidence of such impairment by examining the behavior of financial markets in the United States, the euro area, and New Zealand in light of the communication strategies of the central banks. While we find evidence that central bank policy rate forecasts influence market prices in New Zealand, we find no evidence that market participants in the three regions systematically overweight policy rate guidance or that they do not appreciate the uncertainty and conditionality of it. The results suggest that the risk of impairing market functioning is not a strong argument against central banks' provision of policy rate guidance or forecasts.

JEL Codes: E52, E58, G14.

1. Introduction

When evaluating the advantages and disadvantages of providing to the public guidance about or regular forecasts of policy rates, central bankers have expressed concerns that the provision of such guidance or forecasts may impair financial-market functioning because market participants will place an inordinate amount of weight on them.

*We are grateful to Bill Allen, David Archer, Charlie Bean, Meredith Beechey, Claudio Borio, J. Benson Durham, Michael Ehrmann, Luci Ellis, William English, Marcel Fratzscher, Gabriele Galati, Charles Goodhart, Refet Gürkaynak, Özer Karagedikli, Patricia Mosser, Roberto Perli, Brian Sack, Philip Wooldridge, Jonathan Wright, Christian Upper, participants at seminars at the Austrian National Bank, BIS, ECB, and Sveriges Riksbank, and two anonymous referees for helpful comments and discussions, and to Garry Tang for excellent research assistance. The views expressed are those of the authors and should not be thought to represent those of the BIS or the Federal Reserve Board.

In this paper, we assess the seriousness of these concerns by evaluating the behavior of financial markets in the United States, the euro area, and New Zealand in light of the communication strategies of central banks. While we find evidence that central bank policy rate forecasts influence market prices in New Zealand, we find no evidence that market participants in the three regions systematically overweight policy rate guidance or that they do not appreciate the uncertainty and conditionality of it. The results suggest that the risk of impairing market functioning is not a strong argument against central banks' provision of policy rate guidance or forecasts.

There has been a profound transformation in central bank communication practices over the past two decades. Previously, central banks often shrouded their deliberations, policy intentions, and even policy actions in secrecy. Nearly all central banks now announce publicly their policy actions: most provide detailed information about their policy meetings in the form of minutes, press briefings, or even transcripts; many make their policy intentions clear by announcing inflation targets or other objectives; and some release their economic projections. The current cutting edge of the movement toward greater transparency is the issue of whether or not central banks should provide regular forecasts of their own policy rates. The Reserve Bank of New Zealand (RBNZ) has provided regular forecasts of the ninety-day bank bill rate since June 1997. Among other central banks, those of Norway (in November 2005), Sweden (in February 2007), Iceland (in March 2007), and the Czech Republic (in February 2008) have also begun publishing policy rate forecasts.[1]

Some central banks have opted to provide guidance for finite periods of time about the likely near-term path for policy rates. From April 1999 to August 2000 and from March 2001 to July 2006, the Bank of Japan (BOJ) indicated that its target of zero for the interbank rate would be maintained until deflationary concerns were dispelled. The Federal Open Market Committee (FOMC) signaled the trajectory for rates from August 2003 to December 2005, first by stating that rates would remain at 1 percent for a "considerable period," and then by indicating that the tightening in policy

[1]While the central banks of New Zealand, Norway, and Sweden publish the expected policy rate path of their monetary policy decision makers, the central banks of Iceland and the Czech Republic publish staff forecasts of the policy rate.

would proceed at a pace that was likely to be "measured."[2] The European Central Bank (ECB) telegraphed each of its policy moves during the tightening episode from December 2005 to August 2007 by using language inserted in the statement released following the previous month's policy meeting. "Strong vigilance" always preceded (and was taken by market participants to imply) a tightening at the next meeting, and "close monitoring" preceded unchanged policy. In contrast to the RBNZ, the FOMC and the ECB have not provided explicit or regular forecasts for interest rates.

Little theoretical work has been done so far on whether the provision of forecasts of their own policy rates by central banks is beneficial. Rudebusch and Williams (2006) argue that in an economy where private agents have imperfect information about the determination of monetary policy, central bank communication of interest rate projections is desirable because the projections can help shape financial-market expectations and improve macroeconomic performance. Much of the literature on the effect of central bank transparency has focused on the effects of central bank transparency about exogenous state variables and economic projections, rather than about the policy rate set by the central bank. Svensson (2004) and Woodford (2005) argue that more transparency is better than less, since greater transparency reduces uncertainty about central bank objectives and enhances accountability. Morris and Shin (2002, 2005) argue that market participants will focus too intently on the public forecasts and pay too little attention to other private sources of information. The inattentiveness of market participants to their own private information reduces the information content of market prices.[3]

Central bankers typically see advantages and disadvantages in providing interest rate forecasts (see Archer 2005; Issing 2005; Kohn 2005, 2008; Bergo 2007; Ingves 2007; King 2007; Rosenberg 2007; and Tucker 2007). They recognize the value of reducing uncertainty

[2]In the past, the FOMC provided balance-of-risk assessments suggesting the likely direction of future monetary policy, the impact of which on financial markets has been studied in Ehrmann and Fratzscher (2007b).

[3]Svensson (2006), however, argues that the conclusions of Morris and Shin (2002) depend on implausible parameter assumptions.

about central bank objectives and tactics. They also note that affecting private-sector expectations about future monetary policy is an important means by which central banks influence economic activity. On the other hand, many point out that it can be difficult for monetary policy committees to reach agreement about a forecast for policy rates. For example, Donald Kohn (2005), the Vice Chairman of the Federal Reserve Board, states that "the possibility that discussions of future policy, even nonspecific, could create presumptions about a string of policy actions makes finding a consensus among policymakers on what to say about future interest rates quite difficult—more so than agreeing on the policy today." Goodhart (2001, 172) states, "One alternative would be to have the MPC decide, and vote, not just on the change in interest rates this month but also on the whole prospective path.... The space of choice becomes so great that it is hard to see how a committee could ever reach a majority for any particular time path." Reaching a consensus is not a concern, however, for a sole monetary policy decision maker, as in the case of the RBNZ.

Central bankers also frequently note that central bank forecasts of policy rates run the risk of impairing market functioning. For example, in a speech at the 2005 American Economic Association meetings, Donald Kohn listed two considerations that have constrained the pace of central bank transparency about their outlook: The first consideration is that informational efficiency could be impaired by the provision of policy rate guidance, with financial markets placing too much weight on central bank forecasts, reducing their own analysis of economic developments, and not appreciating sufficiently the uncertainty surrounding these forecasts and their conditionality. The second consideration is the possibility that deviations from policy projections that were too firmly believed by market participants would unsettle financial markets, a possibility that would make it difficult for policymakers to depart from the projected path. For example, Kohn (2005) stated that "in any case, the risks of herding, of overreaction, of too little scope for private assessments of economic developments to show through, would seem to be high for central bank talk about policy interest rates." Issing (2005, 70) stated, "However, with the use of such code words, the central bank puts itself under pressure to honor a quasi-promise. If, in the meantime, its assessment of the situation has changed, owing to new developments, the central bank will be faced with the dilemma of triggering market disturbances if they 'disappoint' expectations, even though they may have convincing arguments to justify their reassessment of the circumstances. For this reason, indications about future decisions must always be seen only as conditional commitments. In practice, however, it is likely to prove extremely difficult to communicate this proviso with sufficient clarity. The more straightforward the 'announcement' and the simpler the code, the more difficult it will be to explain its conditionality ex ante." Goodhart (2001, 175) expressed the concern that "any indication that the MPC is formally indicating a future specific change in rates (e.g., as driven by a 'rule'-based formula) would be taken to indicate some degree of commitment."

In this paper we evaluate these two risks to financial-market functioning about which policymakers have expressed concerns, in light of the communication strategies of central banks. We do so by examining the following four questions: Do policy rate forecasts influence market prices? Are market participants inattentive to other developments when central banks provide policy rate forecasts? Do market participants take policy rate forecasts too seriously? And, do deviations from policy rate forecasts unsettle financial markets? We find evidence that policy rate forecasts do influence market prices, but we find no evidence that the forecasts impair market functioning.

2. Does Policy Rate Guidance Influence Market Interest Rates?

If policy rate guidance does not influence market interest rates, then the guidance would seem unlikely to impair market functioning or, for that matter, be particularly useful. Some studies for the United States have found that policy rate guidance influences U.S. market interest rates. Kohn and Sack (2003) find that statements released by the FOMC significantly affect market interest rates, partly since these statements convey information about the near-term policy inclinations of the FOMC. Gürkaynak, Sack, and Swanson (2005) find that a factor with a structural interpretation as the "future path of policy" significantly influences U.S. market interest rates, with the impact being larger for longer-term U.S. Treasury yields than for shorter-term market interest rates.

While the evidence from the United States is suggestive, the FOMC does not provide an explicit policy rate forecast, unlike the Reserve Bank of New Zealand, which has the longest history of providing forecasts of future policy rates. Therefore, to investigate whether policy guidance influences market rates, we focus on the evidence from New Zealand. The RBNZ has provided forecasts of the ninety-day bank bill rate since June 1997 at various horizons. We use the interest rate projections published by the Reserve Bank of New Zealand in their quarterly Monetary Policy Statements (MPSs), which were published starting in June 1997.[4] The MPSs are published at 9 a.m. New Zealand time on scheduled dates, four times a year. In March 1999, the RBNZ switched from a quantity-based system of implementing monetary policy, which had been accompanied by "open-mouth operations," to a system based on the overnight cash rate (OCR) (see Brookes and Hampton 2000, and Guthrie and Wright 2000). The MPSs are published at the same time as the OCR announcements. In addition, there are four OCR announcements a year not accompanied by an MPS and policy rate forecast, but just by a one-page press release. We match the published interest rate forecasts up to eight quarters ahead with the market interest rates implied by the New Zealand ninety-day bank bill futures contracts, in order to study the relationship of the forecasts with expected future market interest rates.

In order to evaluate the effect of the new central bank interest rate forecast on market interest rates, we would like to evaluate the reaction of the futures rate on the day of publication of the forecast, $(f_n(t) - f_n(t-1))$, to the surprise in the forecast,

$$f_n(t) - f_n(t-1) = c + b(f_n^{CB}(t) - E_{t-1} f_n^{CB}(t)) + \varepsilon_t, \qquad (1)$$

where $f_n^{CB}(t)$ is the central bank's interest rate forecast n quarters ahead made at time t, $f_n(t)$ is the futures rate on the day of publication of the forecast expiring n quarters ahead, $f_n(t-1)$ is the futures

[4]The RBNZ's policy rate forecast is determined endogenously along with inflation and output using their Forecasting and Policy System model (see McCaw and Ranchhod 2002, and Ranchhod 2003).

rate on the day before publication of the forecast, and $E_{t-1}f_n^{CB}(t)$ is the market's expectation of the central bank's forecast on the day prior to its publication.[5]

In the absence of a perfect measure for the market expectation of the central bank's forecast in equation (1), $E_{t-1}f_n^{CB}(t)$, we include two proxy measures for it in the regression. The first proxy is the futures rate on the day prior to publication of the forecast, $E_{t-1}^{(1)}f_n^{CB}(t) = f_n(t-1)$. The second proxy we use is the previous central bank forecast made a quarter ago, $E_{t-1}^{(2)}f_n^{CB}(t) = f_{n+1}^{CB}(t-1q)$. Here, $f_{n+1}^{CB}(t-1q)$ is the forecast $n+1$ quarters ahead made in the previous quarter.[6] The first proxy measure is the most timely one. It should incorporate all the information available to market participants up to the day prior to publication of the central bank's forecasts. However, this measure may contain term premia and therefore may not reflect market participants' expectations accurately. In addition, market participants' true expectations about future interest rates may differ from those of the central bank. We therefore also include the second proxy measure, the central bank's previous forecast, which does not suffer from these two drawbacks and which market participants are likely to factor into their expectations. However, it is a less timely measure and does not include the latest information.

Using these proxies for market expectations of the central bank's forecast, the regression equation for changes in market interest

[5]We consider daily changes in market interest rates in equation (1), rather than intraday changes. Some researchers use intraday data (see, e.g., Andersen et al. 2003). But others use daily data, including Ehrmann and Fratzscher (2004, 2007a). Ehrmann and Fratzscher (2004) argue that intraday data may capture overshooting effects of the market that quickly disappear. Moreover, not all market participants necessarily react to news within a few hours. Based on these arguments, and based on our experience with conducting event studies for the United States and Canada (see Gravelle and Moessner 2002), we use daily data in this study. Drew and Karagedikli (2008) consider the reactions of market interest rates in New Zealand to economic news at both the daily and intraday frequency. They find that daily data give similar results to intraday data for the estimated coefficients, with the coefficients still quite significant for short-term interest rates (which we consider in our paper).

[6]It refers to $n+1$ quarters ahead in order to match the n-quarter-ahead forecast made a quarter later.

rates to surprises in forecasts on the forecast publication dates becomes

$$f_n(t) - f_n(t-1) = c + b(f_n^{CB}(t) - dE_{t-1}^{(1)} f_n^{CB}(t) \\ - (1-d)E_{t-1}^{(2)} f_n^{CB}(t)) + \varepsilon_t, \qquad (2)$$

which we estimate using nonlinear least squares. Table 1 reports the results for these regressions, separately for each horizon n.[7] We can see from table 1 that the surprises in the RBNZ forecasts have a significant influence on financial-market interest rates at horizons of two to six quarters ahead, with coefficients between 0.17 and 0.22. These results are consistent with those reported in Archer (2005), who finds that the New Zealand yield-curve slope is weakly influenced by surprises in the published interest rate slope. On the one hand, these coefficients may appear small, with market interest rates not moving one-for-one with surprises in central bank forecasts. This may suggest that market participants ignore central bank forecasts to a large degree, which may be perceived as damaging the central bank's credibility. On the other hand, we only have imperfect proxy measures available for the market's expectations of the RBNZ forecasts in the regressions, so that their correspondence is not perfect, and coefficients below 1 would be expected due to this measurement problem. Moreover, no doubt at least to some extent the central bank forecast is surprising to market participants because the central bank has changed its views about the likely future path for interest rates for reasons that market participants do not find compelling, and so a coefficient below 1 should be expected.[8]

[7] Another alternative is to regress the central bank forecast on the two proxies for the market's expectation and use the residual from the regression as a measure of the surprise component of the forecast. Using this alternative approach yields very similar results for the coefficients on the surprises reported in table 1. We prefer the specification reported in table 1, however, since it does not use a derived measure for the surprise in the regression.

[8] As discussed above, the futures rates will also not equal expected future interest rates because of term premia. However, term premia might be expected to be fairly small at the horizons we consider.

Table 1. Reaction of Daily Changes in Interest Rate Futures to Surprises in RBNZ Interest Rate Forecasts on the Days of Publication of the Forecasts

Quarters Ahead	**1**	**2**	**3**	**4**	**5**	**6**
Constant, c	0.01	−0.005	0.001	−0.01	−0.01	−0.02
	(0.4)	(−0.2)	(0.0)	(−0.8)	(−0.6)	(−0.7)
Surprise in Forecast, b	0.13	0.20**	0.20**	0.22**	0.20**	0.17**
	(1.9)	(3.6)	(4.6)	(5.9)	(5.5)	(4.1)
First Proxy for Expected Forecast, d	0.00	0.51**	0.43**	0.52**	0.44**	0.45**
	(0.0)	(2.8)	(2.9)	(4.7)	(3.8)	(2.8)
No. of Observations	39	39	39	33	30	29
R^2	0.20	0.29	0.38	0.54	0.53	0.40
LM Test for Serial Correlation of	1.30	1.56	1.05	0.19	0.53	0.90
Residuals[1]	[0.29][2]	[0.21][2]	[0.40][2]	[0.94][2]	[0.72][2]	[0.48][2]

Notes: t-values are in parentheses; * and ** denote significance at the 5 percent and 1 percent level, respectively.
[1]Breusch-Godfrey LM test with four lags, F-statistic.
[2]p-values are in square brackets.
The first proxy for the expected forecast, with weight d, is the New Zealand ninety-day bank bill futures rate on the day prior to publication of the forecast, the same number of quarters ahead as the forecast; the second proxy for the expected forecast, with weight $1 - d$, is the previous central bank forecast made a quarter ago. The sample is from June 27, 1997, to March 8, 2007, at quarterly intervals on the dates of publication of the interest rate forecasts in the RBNZ's Monetary Policy Statements. The New Zealand ninety-day bank bill futures contracts are traded on the Sydney futures exchange.

3. Are Market Participants Inattentive to Other Developments When Central Banks Provide Policy Rate Guidance?

A common concern raised by central bankers is that market participants may pay too much attention to policy rate forecasts and pay too little attention to other sources of macroeconomic information. As a consequence, if policy rate forecasts are provided, market prices would become less informative. To investigate this possibility, we examine the response of interest rate futures and option-implied volatilities to macroeconomic data releases and central bank policy announcements. We find no evidence that policy guidance leads market participants to reduce their reaction to other sources of news.

3.1 Response of Interest Rate Futures to Economic Data Releases

One might expect that during the period when the FOMC was providing clear signals about future monetary policy (August 2003 to December 2005), the sensitivity of asset prices to macroeconomic releases might fall, insofar as the FOMC was signaling that future policy adjustments would be gradual. However, we find that the responsiveness of one-year-ahead Eurodollar futures rates to a set of major macroeconomic releases was significantly higher during the guidance period (see table 2). Table 2 reports results for the regressions of daily changes in one-year-ahead Eurodollar futures rates (in basis points), $y(t) - y(t-1)$, on the surprise components of eleven economic releases,

$$
\begin{aligned}
y(t) - y(t-1) = c + c_g \; dum_g(t) + \sum_{e=1}^{11} (b_e \; surprise_e(t) \\
+ g \; b_e \; surprise_e(t) \; dum_g(t)) + \varepsilon_t, \qquad (3)
\end{aligned}
$$

where the subscript e denotes changes in nonfarm payrolls, the unemployment rate, hourly earnings, CPI inflation, PPI inflation, industrial production, the trade balance, retail sales, housing starts,

Table 2. Difference in the Effect of Macroeconomic Data Releases on Daily Changes in One-Year-Ahead Eurodollar Futures Rates When the FOMC Was Providing Rate Guidance[a]

Variable	Estimate
Constant	−0.5**
	(−2.9)
Guidance Dummy on Constant (c_g)	0.5
	(1.6)
Nonfarm Payrolls	5.0**
	(7.8)
Unemployment	−1.9**
	(−4.4)
Hourly Earnings	1.0*
	(2.5)
CPI	0.5
	(1.4)
PPI	−0.3
	(−0.7)
Industrial Production	0.3
	(1.0)
Trade Balance	0.5
	(1.4)
Retail Sales	1.8**
	(3.8)
Housing Starts	−0.3
	(−0.8)
ISM	1.1**
	(2.9)
GDP	−0.1
	(−0.2)
Common Proportional Change in Response during Guidance Period (g)	2.1**
	(4.4)
No. of Observations	2,247
R^2	0.10

Notes: * and ** denote significance at the 5 percent and 1 percent level, respectively. t-values are in parentheses. [a]Daily changes in basis points; the sample period is from June 1998 to August 2007. The guidance period is defined in the note to table 3. The macroeconomic data surprises are calculated relative to the median of the most recent Bloomberg survey and are normalized by their standard deviation.

the ISM manufacturing index, and GDP.[9] The surprises are calculated relative to Bloomberg median survey expectations and are normalized by their standard deviation. The guidance dummy, $dum_g(t)$, is equal to 1 during periods when the FOMC provided guidance and 0 at all other times. The coefficient b_e is the estimated response, in basis points, of one-year-ahead Eurodollar futures rates to a one-standard-deviation surprise in the economic statistic outside the guidance period, and $(1 + g)b_e$ is the response during the guidance period. A significantly negative estimate of g would indicate reduced responsiveness during the guidance period, while a significantly positive estimate would indicate increased responsiveness.

We can see from table 2 that the coefficients on five of the economic releases are statistically significant and of the expected sign. The coefficient g is estimated to be 2.1 and is highly significant, indicating that the response of interest rate futures was significantly stronger during the guidance period. These results suggest that financial-market participants continued to pay close attention to macroeconomic information during the period when the FOMC was providing guidance on future policy rates. Our finding that the reaction to macroeconomic surprises does not decrease significantly during the guidance period is consistent with a result of Ehrmann and Fratzscher (2007b), who find only weak evidence for a reduction in market reactions to macroeconomic surprises with the introduction of balance-of-risk assessments by the FOMC in 1999.[10]

3.2 Response of Interest Rate Futures to Policy Announcements

If market participants shift their focus toward policy announcements when policy rate guidance is provided, the responsiveness of the outlook for future interest rates to monetary policy releases should increase *relative* to the response to other sources of information. To

[9]The releases are essentially the same as those considered by Gravelle and Moessner (2002).

[10]By contrast, there is some evidence that greater transparency in the form of publication of an inflation target can lead to a better anchoring of private agents' long-term inflation expectations and reduce the sensitivity of long-term inflation expectations derived from government bond yields to economic news (see Gürkaynak, Levin, and Swanson 2006, and Libich 2006).

examine that hypothesis, we consider the ratio of the absolute values of daily changes in one-year-ahead interest rate futures on monetary policy announcement days to the averages of those changes over recent periods (up to N days previously),

$$r_i^a(t) = 100|y_t - y_{t-1}|^i / \sum_{n=1}^{N} (|y_{t-n} - y_{t-n-1}|^i / N), i = 1 \text{ or } 2, \quad (4)$$

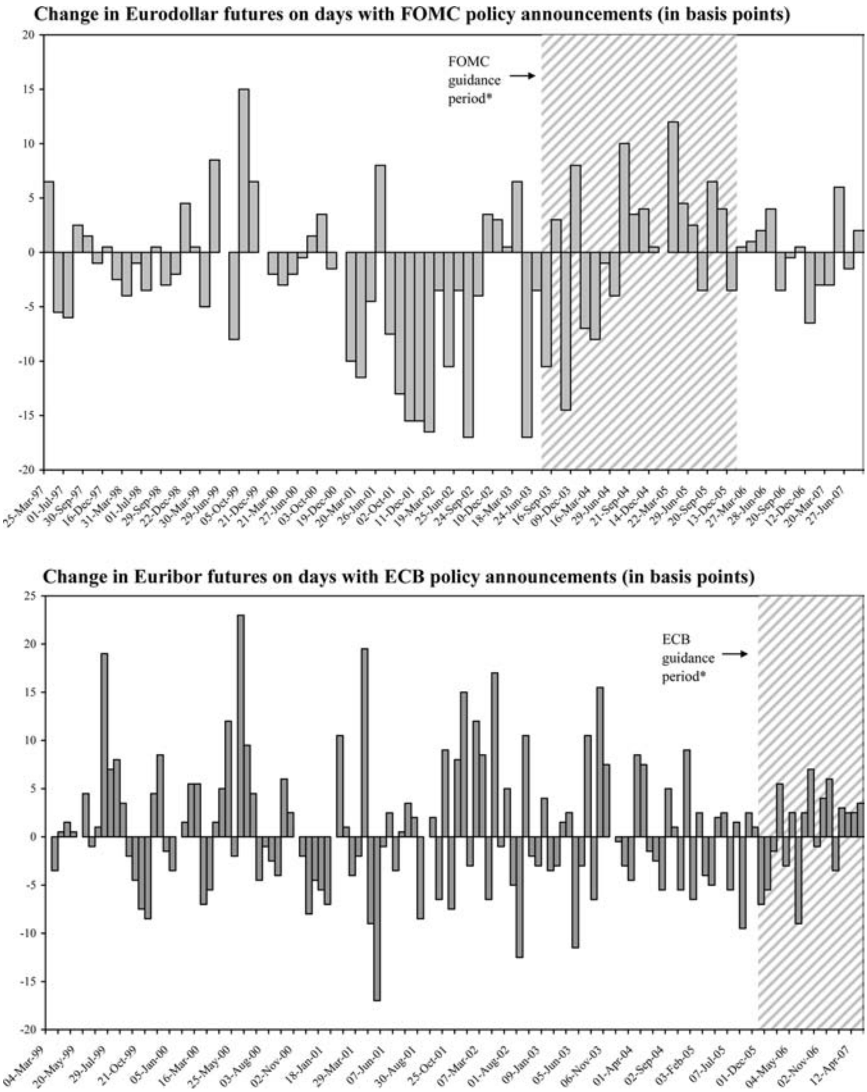
on the view that movements on contiguous non-policy-announcement days will reflect the responsiveness of rates to other sources of information. We then compare the ratio over all the monetary policy days in our sample with those policy days when the central banks were providing guidance about the future path of interest rates. We look at both the FOMC and the ECB. In general, we find no evidence that there is a significant increase in this ratio during the periods with guidance.

The one-year horizon is short enough that the interest rate futures are determined primarily by expectations about monetary policy but long enough not to be nailed down by any implicit commitment to specific monetary policy choices inherent in the central banks' statements about the outlook for policy rates. During these episodes, the central bank communications were designed to telegraph near-term policy choices, and there were virtually no surprises in the precise choice of policy rates at each meeting. Judging by the changes in one-year-ahead interest rate futures, however, during the guidance periods, revisions to the outlook for the path of policy beyond the very near term were just about as volatile as during other periods (see figure 1).

For the FOMC, we use the absolute value of daily changes in one-year-ahead Eurodollar futures as our measure of the revision to the interest rate outlook. The change on the day of FOMC meetings is divided by the average change over the preceding four weeks (the FOMC meets about every six weeks). The sample begins in 1994, when the FOMC first began releasing press statements when it changed policy. The results for the following regressions are shown in table 3,

$$r^a(t) = c + b\ dum_g(t) + \varepsilon_t, \quad (5)$$

Figure 1. Changes in Eurodollar and Euribor Futures on Days with Policy Announcements



where again the guidance dummy, $dum_g(t)$, is equal to 1 during periods when the FOMC provided guidance and 0 at all other times. The absolute changes on all FOMC days average 33 percent higher than other days over the preceding month.

Table 3. Ratio of Absolute Value of Changes in Interest Rate Futures on Policy Announcement Days to Other Days during Periods When Central Banks Provide Policy Guidance

	Federal Reserve	**ECB**
Constant, c	132.7**	157.95**
	(11.6)	(10.6)
Guidance Dummy, b	8.4	−6.8
	(0.3)	(0.2)
No. of Observations	109	64
R^2	0.00	0.03

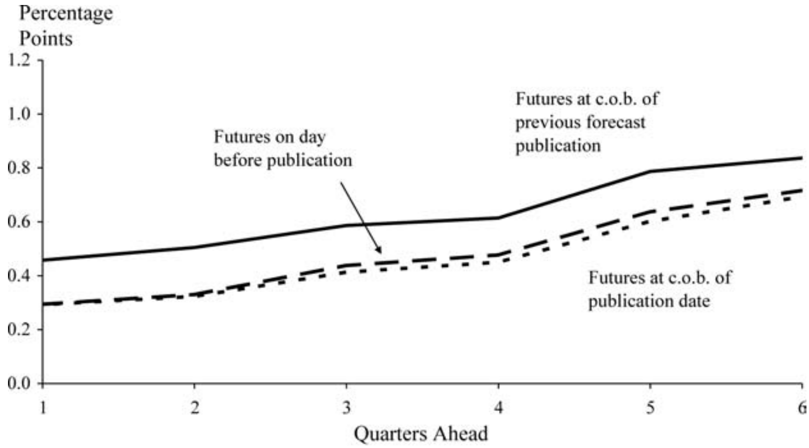
Notes: * and ** denote significance at the 5 percent and 1 percent level, respectively. t-values are in parentheses. The FOMC provided guidance about the likely trajectory for policy from August 12, 2003, to December 13, 2005, when it first indicated that interest rates would be held at 1 percent for a "considerable period" and then stated that policy tightening would proceed at a pace likely to be "measured." The sample consists of the 109 FOMC meetings from February 1994 to August 2007. The ECB telegraphed its policy moves one month in advance from December 2005 to August 2007. The sample consists of the sixty-four monetary policy announcements from January 2002 to August 2007.

The changes are an additional 8 percentage points higher on meeting days when the FOMC was using the "considerable period" and "measured pace" language, but the difference is not statistically significant.

For the ECB, we use the daily changes in one-year Euribor futures, and the sample begins in January 2002. We only use a three-week moving average as the denominator so that the period does not include a previous meeting day (the Governing Council of the ECB has met once a month to decide on the policy rate since 2002).[11] We test to see if the relative variance on meeting days rose during the period when the ECB's President Trichet alternated between

[11] We start in 2002 since prior to 2002, the ECB's Governing Council met twice monthly at scheduled meetings to decide on monetary policy, although policy rates were generally not changed at the meeting in the middle of the month. In September 2001, the Governing Council met three times for monetary policy decisions.

Figure 2. Mean Absolute Difference between Futures[a] and RBNZ Forecasts[b]



[a] Futures on New Zealand ninety-day bank bills.
[b] Forecasts by Reserve Bank of New Zealand of ninety-day bank bill rates.

"strong vigilance" and "close monitoring" to signal if the next move would be a 25-basis-point increase or no change, respectively.

The results provide even less evidence that markets became overly attentive to the ECB's policy announcements or press conferences. For the sample as a whole, the absolute value of the interest rate changes is 58 percent higher on policy announcement days than on other days during the preceding three weeks. The boost on meeting days is 7 percentage points *less* during the signaling period, but the decrease is not statistically significant.

In New Zealand, market participants' and the RBNZ's forecasts of the ninety-day bill rate have moved together closely over time. This result is illustrated in figure 2, separately for each horizon n quarters ahead, for the forecasts published between June 1997 and March 2007. Figure 2 shows the mean absolute difference between the published forecast and the futures rate on the day of publication of the forecast, the futures rate on the day prior to publication, and the futures rate the day the previous forecast was published. While the futures rate moves closer to the forecast on the day the forecast is published, that narrowing of the gap is small compared with the narrowing that occurs over the quarter up to the day prior to the forecast. The narrowing of the difference between the futures rates and the forecasts occurring over the quarter up to the day prior to the release of the forecast cannot, of course, reflect a response to the as-yet-unknown RBNZ forecast. This suggests that both forecasts and futures are to some extent reacting to the same news about the economic outlook arriving between forecast publication dates, to the OCR announcement and accompanying press release occurring in between the publication of the MPSs, or that additional changes in the RBNZ's policy outlook are revealed to the market in speeches, testimonies, or by other means. That is, futures rates adjust to new information arriving between the publication of forecasts, and market participants do not just react to published forecasts. The RBNZ forecasts may also be influenced by movements in interest rate futures.

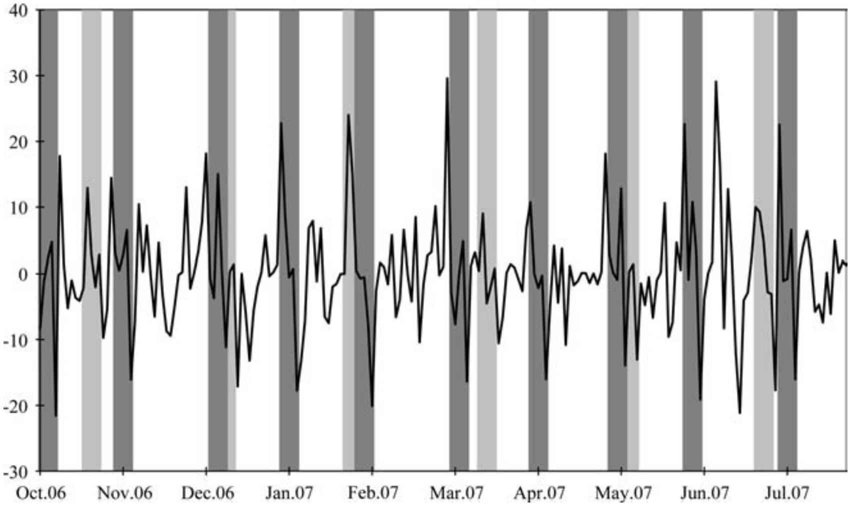
3.3 Response of Option-Implied Interest Rate Volatility

We can also evaluate the relative impact on market interest rates of central bank policy announcements by examining the behavior of the option-implied volatility of interest rates. The implied volatilities we use are taken daily from over-the-counter options on five-year Treasury securities with one week to maturity.[12] We examine the implied volatility in five-year yields rather than at shorter maturities because we do not have data available on shorter-maturity securities for options with a constant, short period to expiry. While movements in five-year-ahead *futures* or *forward* rates may have little to do with monetary policy, the five-year Treasury yield is a yield to maturity and not a forward rate, and so is significantly influenced by interest rates expected for the next few years, which are determined importantly by monetary policy expectations. For example, the correlation between the daily changes in five-year and two-year Treasury yields was 0.92 over the period 1990 to mid-2007.

These data are different from the implied volatility data most commonly used. Usually, daily time series on implied volatility are taken from a single option or interpolated from two options with constant maturity dates. Once a quarter, the reference options switch to

[12] Goldman Sachs has generously provided us with these data.

Figure 3. Option-Implied Volatility of Five-Year U.S. Treasury Note Yields, Effect of Employment Reports and FOMC Announcements



Note: Daily. The implied volatilities are from options with one week to maturity. The shaded regions are the one-week periods prior to FOMC announcements (light grey) and U.S. employment reports (dark grey).

ones maturing one quarter later. The maturity dates are often several quarters into the future. The data we use here measure implied volatility using a different option each day. The option chosen always expires in one week. The short and constant maturity of the options allows us to measure the increase in implied volatility when specific events enter the relevant window and the decline when the events leave the window.[13] As can be seen in figure 3, when specific risk events—in this case, employment reports or FOMC meetings—are scheduled to occur within the week remaining till the option expires, the volatility in five-year yields over the week implied by the option price is noticeably higher.

Since these data are only available to us for the United States, we can only test for the impact of U.S. economic news and FOMC announcements, not for similar events in other countries. We look

[13] We benefited from discussions with Brian Sack concerning this procedure.

at the effects of FOMC announcements and of macroeconomic releases from March 1994 to the present, and test for any difference in the effects during the period when the FOMC was providing policy outlook guidance. Because of risk premia, implied volatilities are only imperfect proxies for market participants' true uncertainty. However, daily variations in such risk premia are likely to be small, so that they would not be expected to significantly affect regression results involving daily changes in implied volatilities.

Specifically, we regress the daily log-difference in the implied volatility, $100^*(log(iv(t)) - log(iv(t-1)))$, on one or more sets of two dummies. The first dummy, the event dummy $dum_e(t)$, is equal to 1 on the day one week before the event of interest (when the event can first begin to influence the payoff of the underlying options) and equal to -1 on the day of the event (when the potential influence ends).[14] The second dummy is the event dummy interacted with a variable that equals 1 during the period when the FOMC was providing guidance ($dum_g(t)$). When analyzing the effect of FOMC meetings, the regression is of the form

$$\begin{aligned} 100^*(log(iv(t)) - log(iv(t-1))) &= c + b\ dum_e(t) \\ &\quad + d\ dum_e(t)\ dum_g(t) + \varepsilon_t, \quad (6) \end{aligned}$$

where the subscript e denotes FOMC meetings. When analyzing the effect of data releases, we use an approach similar to the one we report above for measuring the effect of data releases on interest rate futures. The daily percentage change in volatility is regressed on the dummies for each of the data releases described above as well as those dummies interacted with the "guidance period" dummies. A coefficient is estimated for each data release and a single additional coefficient measures the proportional change in the effect of all the data releases during the guidance period. The data releases considered are the same set as considered above—namely, the employment

[14]We are restricting the increases and decreases in implied volatility to be equal, a restriction accepted by the data.

report (which includes changes in nonfarm payrolls, the unemployment rate, and hourly earnings released at the same time), CPI inflation, PPI inflation, industrial production, the trade balance, housing starts, retail sales, the ISM manufacturing index, and GDP. In this case, the equation takes the form

$$100^{*}(log(iv(t)) - log(iv(t-1))) = c + \sum_{e=1}^{9}(b_e\ dum_e(t) + g\ b_e\ dum_e(t)\ dum_g(t)) + \varepsilon_t, \tag{7}$$

where the subscript e denotes the data releases. The release increases implied volatility by b_e percent outside of the guidance period and $(1 + g)b_e$ percent during the guidance period. The results are reported in table 4.

The implied volatility of the five-year U.S. Treasury yield behaves in a manner consistent with the results for the ex post volatility of interest rate futures discussed above. Implied volatilities are higher by about 5 percent when an FOMC meeting occurs during the week before the option expires, and the increase is highly statistically significant. The increase was 0.3 percentage point greater during the "considerable period/measured pace" interval, but the difference is not statistically significant.

The presence of an economic statistical release during the week before the option expires also generally boosts implied volatilities. Seven of the nine releases considered increased implied volatilities by a statistically significant amount. Employment reports were, in fact, viewed as more consequential risk events for five-year yields than FOMC meetings, increasing implied volatilities by 18 percent (see table 4 and figure 3). The impact of economic data releases on implied volatilities did not fall during the guidance period; it rose by a highly statistically significant 70 percent.

In sum, the behavior of implied volatilities provides no evidence that the FOMC's guidance led market participants to be inattentive to other sources of information. FOMC announcements were expected to have about the same impact on five-year yields as during other times, and macroeconomic releases were expected to have, if anything, a larger impact.

Table 4. Percent Increase in the Option-Implied Volatility Caused by FOMC Meetings and Macroeconomic Releases during Periods When the FOMC Provided Policy Guidance

Variable	Estimate
Constant	0.0
	(0.1)
FOMC Meeting Dummy	5.4**
	(6.5)
FOMC Meeting Dummy*Guidance	0.3
Period Dummy	(0.1)
No. of Observations	3,221
R^2	0.02
Constant	−0.0
	(−0.2)
Employment Report	17.8**
	(32.0)
CPI	3.5**
	(7.5)
PPI	2.2**
	(4.8)
Industrial Production	−0.2
	(−0.3)
Trade Balance	−0.3
	(−0.6)
Retail Sales	1.9**
	(4.2)
Housing Starts	1.4**
	(3.0)
ISM	1.9**
	(4.4)
GDP	2.5**
	(3.3)
Common Proportional Change in	0.7**
Coefficients during Guidance Period (g)	(7.8)
No. of Observations	3,221
R^2	0.3

Notes: * and ** denote significance at the 5 percent and 1 percent level, respectively. t-values are in parentheses. The sample is from March 31, 1994, to July 27, 2007. The guidance period is defined in the note to table 3. The implied volatilities are taken from options with one week to expiration on five-year Treasury notes. The regressions estimate the increase and decrease (restricted to be equal) in the implied volatility when FOMC meetings, or the indicated economic releases, enter and leave the one-week window.

4. Do Market Participants Take Central Bank Policy Rate Guidance Too Seriously?

A slightly different concern commonly raised by central bankers is that market participants will not understand that central banks' statements about future policy rates are not commitments, that the statements are conditional on developments or are forecasts subject to uncertainty and error. Put another way, when provided with forecasts or policy guidance, do market participants become excessively confident in their outlook for interest rates?

When evaluating this concern, it seems important to distinguish between providing forecasts on a regular basis, such as is done by the Reserve Bank of New Zealand, and including forward-looking language in policy announcements for a temporary period. When a forecast is released regularly, the central bank is not making a tactical decision to release or not release the forecast, and so the *existence* of the forecast does not necessarily imply anything about the central banks' intentions (which is not to say that the content of the forecast does not convey information about the central banks' intentions).[15] However, if a central bank only sometimes provides guidance about future policy, the central bank is making a tactical decision to manage market participants' expectations. As a result, when a central bank only sometimes provides guidance, the *existence* of the guidance may imply some degree of commitment and therefore a lower level of uncertainty about near-term policy rates.

For example, the minutes of the August 2003 FOMC meeting indicate that the FOMC foresaw keeping policy accommodative for a "considerable period" because it was concerned about the risk of deflation and anticipated keeping interest rates lower than normal in the future when the economy strengthened.[16] Similarly, when the FOMC adopted the "measured" language in May 2004, it indicated that the tightening would likely be more gradual than normal

[15]However, Rosenberg (deputy governor of the Sveriges Riksbank) mentioned in a speech that one of the motivations for a central bank to publish forecasts of its policy rate was to steer expectations (see Rosenberg 2007).

[16]Minutes of the Federal Open Market Committee, August 12, 2003.

because inflation was so low.[17] We do not have minutes for the ECB's meetings and so cannot determine the thinking behind its strategy to signal its tightening moves one meeting in advance. Nor did President Trichet explain the reasons for providing guidance in his press conferences over the period. It seems likely, however, that the FOMC and the ECB both chose to provide fairly explicit guidance when they were beginning a tightening episode in order to prevent long rates from rising sharply, imparting too large a degree of financial restraint.

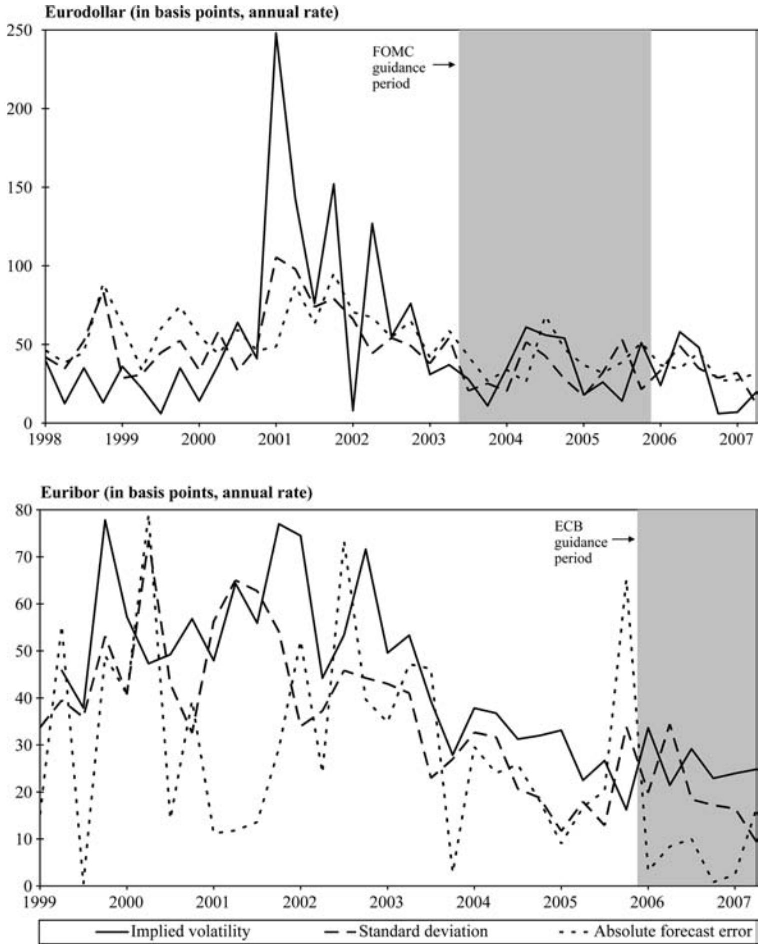
In sum, the issue is not whether market participants take central bank statements about future policy as involving some degree of commitment, but whether they take the statements *too* seriously. To evaluate this possibility, we compare investors' assessments of the uncertainty in the policy outlook as measured by implied volatilities with realized volatilities or forecast errors. If realized volatilities or forecast errors are larger relative to implied volatilities during periods when central banks provide guidance, then the guidance is possibly being taken too seriously.

We look at option-implied volatilities from futures contracts on money-market interest rates with three months to expiration for the United States and the euro area. Implied volatilities are derived from option prices under the assumption that the reference price evolves according to geometric Brownian motion. Under this assumption, it is reasonable to compare implied volatilities with the standard deviation of the daily changes in interest rates. Brownian motion is not, however, a particularly good assumption for interest rate futures prices, since interest rate changes are serially correlated and are subject to jumps, so we also compare the implied volatilities with the realized errors. Neither procedure suggests that market participants are unduly confident about monetary policy when forward-looking guidance is provided by the central bank.

Figure 4 presents the implied volatilities, standard deviations of interest rate changes, and absolute values of the forecast errors for the United States and the euro area. For the United States, the option-implied volatility is for Eurodollar futures with three months to expiration, and the realized standard deviation is for the

[17] Minutes of the Federal Open Market Committee, May 4, 2004.

Figure 4. Implied Volatilities, Standard Deviations of Interest Rate Changes, and Absolute Values of Forecast Errors for Eurodollar and Euribor Futures



daily first difference in the underlying Eurodollar futures rate.[18] The errors are calculated as the difference between the futures rate with three months to expiration (at the same time as the measurement of

[18]The implied volatility is the normalized "basis point" volatility, not the "interest rate" volatility, and so measures the uncertainty in absolute terms around the expected rate, not as a percentage of that rate.

implied volatility) and the spot rate at settlement. We are assuming that term premia will have a negligible, or at least a constant, effect on the realized error over the three-month horizon. For the euro area, the implied volatilities are from Euribor futures, the standard deviation from Euribor futures rates, and the forecast errors are calculated using the Euribor futures and spot Euribor rate. The dummy is defined for the "vigilance/monitoring" interval.

As can be seen, the implied volatilities, especially in the United States, fell to particularly low levels during the period when the central banks were providing interest rate guidance. However, the investor confidence appeared to be warranted, as the realized standard deviations and forecast errors were also quite low. Indeed, Swanson (2004) finds a downward trend in implied volatility that he attributes to investors' ability to forecast interest rates and to increased FOMC transparency.

The impressions from figure 4 are confirmed by regression results reported in table 5. We regress the ratio of the realized standard deviation, $s(t)$, to the implied volatility, $iv(t)$, or the ratio of the absolute value of the realized forecast errors, $fe(t)$, to the implied volatility on a constant and a dummy for the "considerable period/measured pace" interval for the United States and the "vigilance/monitoring" interval for the euro area,

$$r_i^b(t) = c_i + b_i \; dum_g(t) + \varepsilon_t, \; i = 1 \text{ or } 2, \tag{8}$$

where $r_1^b(t) = s(t)/iv(t)$ and $r_2^b(t) = fe(t)/iv(t)$.[19] In both the United States and the euro area, the standard deviations were a touch *lower* relative to the implied volatilities during the interval when guidance was provided, but in neither case were the differences statistically significant. The forecast errors were a bit higher in the United States and a bit lower in Europe relative to the implied volatilities during the relevant periods, but again, neither result is statistically significant.

Market participants have been very confident in their outlooks for near-term money-market interest rates when central banks have provided guidance. Judging by the muted changes in interest rates

[19]More details on the exact definitions of these variables are given in the notes to table 5.

Table 5. Ratio of Realized Money-Market Rate Volatility to Implied Volatility during Periods When Central Banks Provide Policy Guidance

	Standard Deviation of Daily Changes in Futures Rate ($i = 1$)	Absolute Value of Forecast Error ($i = 2$)
Eurodollar, FOMC Guidance		
Constant, c_i	0.9**	0.8**
	(15.1)	(6.3)
Guidance Dummy, b_i	−0.1	0.1
	(0.4)	(0.3)
No. of Observations	50	50
R^2	0.00	0.00
Euribor, ECB Guidance		
Constant, c_i	0.8**	0.8**
	(12.1)	(5.8)
Guidance Dummy, b_i	−0.1	−0.6
	(0.5)	(1.6)
No. of Observations	33	33
R^2	0.00	0.08

Notes: * and ** denote significance at the 5 percent and 1 percent level, respectively. t-values are in parentheses. The guidance periods are defined in the note to table 3. The implied volatility is measured three months before the maturity of Eurodollar futures contracts. The standard deviations are for the daily first differences of the futures rates for the three months before maturity. The forecast error is the expected Eurodollar rate implied by futures prices three months prior to maturity minus the spot rate at maturity. All variables are in basis points and are at an annual rate. The quarterly observations run from 1995:Q1 to 2007:Q2. The Euribor measures of uncertainty are defined in the same way as the Eurodollar measures and extend from 1999:Q2 to 2007:Q2.

and the accuracy of interest rate forecasts embedded in futures rates, however, that confidence was justified and did not reflect a tendency for investors to take the guidance too seriously. Of course, it remains possible that the stability and predictability of realized rates during the guidance periods are the consequence of the monetary policymakers assiduously avoiding surprises after their guidance was misconstrued, but we find this possibility implausible in part, as discussed in the next section, because there is no evidence that monetary policy surprises have had notably bad effects on financial markets.

We also see no evidence of an overreaction in financial markets to surprises in the Reserve Bank of New Zealand's interest rate forecasts. As shown in table 1, surprises in central bank forecasts lead to some reaction of market interest rates, but with a coefficient much less than 1 at all horizons. Moreover, as mentioned above, Archer (2005) finds that the market yield-curve slope in New Zealand is only weakly influenced by surprises in the published interest rate slope, which also suggests no overreaction by financial markets to surprises in the central bank's interest rate forecasts.

Forecasts in which the policy rate projection is endogenous, as in the case of the RBNZ, are likely to be more credible than those made under the assumption of an exogenous interest rate path. Hence the conclusions regarding the effects of publishing policy rate forecasts may differ in the two cases. Moreover, in countries without an explicit inflation target, the publication of policy rate forecasts may have greater informational content than in countries with an explicit inflation target, since it may contain some signal about the level of an implicit target. Geraats (2005) shows that publishing economic forecasts may send a signal about the inflationary intentions of a central bank. Considering the case of New Zealand, which has an explicit inflation target, may therefore understate the importance of publishing policy rate forecasts for countries such as the United States and the euro area without an explicit inflation target.

5. Do Deviations from Earlier Policy Guidance Unsettle Markets?

The final concern raised by central bankers that we consider is that, insofar as market participants take policy rate guidance too seriously, deviations from the foreshadowed policy paths will unsettle markets. The concern is not only that the stability of financial markets and institutions will be lessened, but also that policymakers' awareness of the potential consequences of deviations will constrain their future decisions.[20]

Undoubtedly, central bankers are concerned about the consequences for financial stability of surprising markets. For instance, when the FOMC began tightening policy in 1994 after a long pause, Chairman Alan Greenspan argued that, even though he believed that there was a case on macroeconomic grounds for a 50-basis-point tightening, the first tightening should be only 25 basis points because the surprise would rattle financial markets.[21] If central bankers believed that their statements about the future would be taken too seriously, then those statements would increase the expected magnitude of monetary policy surprises and so could reasonably be seen as posing a risk of constraining future policy choices.

As shown above, however, we find no empirical evidence that policy guidance is, in fact, taken too seriously. In this section, we test if monetary policy surprises are more likely to lessen financial stability during periods when central banks are providing guidance. Unsurprisingly, we find no compelling evidence of such increased sensitivity.

We assess financial stability using option-implied volatilities, $iv(t)$, of ten-year Treasury yields and of the S&P 500 stock index. We measure monetary policy surprises, $mps(t)$, using two variables. The first—the target surprise, $ts(t)$—is the absolute value of the difference on days of FOMC monetary policy announcements between the FOMC's target federal funds rate and the target expected on the eve of the announcement, judging by federal funds futures rates. The second measure—the path surprise, $ps(t)$—is the absolute value of the change in the one-year-ahead Eurodollar futures rate on announcement days.[22] We interact each surprise measure with a dummy,

[20]See Kohn (2005).

[21]FOMC transcript, February 3, 1994, p. 55. In the event, there were a number of abrupt jumps in longer-term interest rates, and associated market volatility, during the tightening episode that began in 1994 as the market and the FOMC reacted to economic data that came in stronger than expected. It would seem reasonable to suppose that the 1994 experience contributed to the decision by the FOMC to use the "measured pace" language during the tightening episode that began in 2004.

[22]We are following the "target" and "path" surprise terminology of Gürkaynack, Sack, and Swanson (2005), but we do not calculate the path surprise

$dum_g(t)$, equal to 1 for the "considerable period/measured pace" interval,

$$100^*(log(iv(t)) - log(iv(t-1))) = c + a\ ts(t) + b\ ps(t) \\ + d\ dum_g(t)\ ts(t) + f\ dum_g(t)\ ps(t) + \varepsilon_t. \quad (9)$$

The significance of the coefficient on the interaction term measures the effect of the guidance on market sensitivity to monetary policy surprises.

The results are reported in table 6. The monetary policy surprise measures do not have a significant effect on these measures of financial stability. In no case was the effect of the policy surprises significantly different during the "considerable period/measured pace" FOMC announcements. By these measures, monetary policy surprises generally do not unsettle markets, and they did not unsettle markets by more when the FOMC was providing forward guidance.[23]

As noted earlier, however, target surprises were very low during the guidance period because the guidance (and other communications) left little doubt about the policy outcome for each meeting. In this sense, we are not strictly testing the effect of departing from a past policy rate forecast. Still, the *path* surprises were substantial during the period (see figure 1). Revisions to the expected path in response to the new statements are an indication that the FOMC was seen as likely to follow a path somewhat at odds with the path market participants previously expected.

The results for New Zealand presented in table 1 above suggest that surprises in central bank interest rate forecasts influence

---

as a residual from a regression of the change in the futures rate on the target surprise.

[23]We also evaluated the effect of monetary surprises on two measures of financial stress developed at the Federal Reserve Board—a broad index of financial stress and the odds that multiple financial institutions will default over the subsequent year (see Nelson and Perli 2005). The results are the same as those reported above using implied volatilities. The effects of the policy surprises are not significantly different during the "considerable period/measured pace" FOMC announcements for either measure. We also find no significant effect when, instead of the absolute value of the surprises, we use the level of the surprises or the squared surprises.

Table 6. Effect on Measures of Financial Stability of Monetary Policy Surprises When the FOMC Provided Policy Guidance

	Percentage Change on Day of FOMC Meeting	
	Implied Volatility of Ten-Year Treasury Note	**Implied Volatility of S&P 500 Index**
Constant, c	−2.1**	−2.9**
	(3.8)	(3.0)
Target Surprise	−1.0	11.2
	(0.2)	(1.1)
Path Surprise	9.3	0.3
	(1.4)	(0.0)
Target Surprise*	−34.8	−29.4
Guidance Dummy	(0.5)	(0.3)
Path Surprise*	−4.9	12.1
Guidance Dummy	(0.4)	(0.6)
No. of Observations	97	97
R^2	0.03	0.02

Notes: * and ** denote significance at the 5 percent and 1 percent level, respectively. t-values are in parentheses. The "target surprise" is the absolute value of the difference between the target for the federal funds rate announced by the FOMC and the target expected on the eve of the announcement as implied by federal funds futures rates. The "path surprise" is the absolute value of the change in the one-year-ahead Eurodollar futures rate on the day of the announcement. The guidance period is defined in the note to table 3. The sample consists of the ninety-seven days with FOMC announcements from August 22, 1995, to May 9, 2007.

changes in market interest rates, with coefficients below 1, suggesting that there is no disorderly overreaction to unexpected changes in projected interest rates and no unsettling effect on financial markets.

6. Conclusions

Central bank communication has changed dramatically over the past decade, with some central banks providing guidance about, or explicit forecasts of, likely future policy rates. One frequently made argument against the provision of such guidance or forecasts is that it runs the risk of impairing market functioning. In this paper, we evaluated the behavior of financial markets in the United States, the euro area, and New Zealand in light of the communication strategies of central banks. We found evidence for New Zealand that central bank forecasts of policy rates influence market prices, but we found no evidence that forecasts or guidance impair market functioning in the United States, the euro area, or New Zealand. In particular, market participants do not appear inattentive to other developments when central banks provide policy rate guidance; they do not appear to take central bank policy rate guidance too seriously; and deviations from earlier policy guidance do not appear to unsettle markets. Consequently, this evidence suggests that concerns about impairing market functioning are not a strong argument against central banks' provision of policy rate guidance or forecasts.

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega. 2003. "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange." *American Economic Review* 93 (1): 38–62.

Archer, D. 2005. "Central Bank Communication and the Publication of Interest Rate Projections." Mimeo, Bank for International Settlements.

Bergo, J. 2007. "Interest Rate Projections in Theory and Practice." Speech at the Foreign Exchange Seminar of the Association of Norwegian Economists, Sanderstølen, January 26.

Brookes, A., and T. Hampton. 2000. "The Official Cash Rate One Year On." *Reserve Bank Bulletin* 63 (2): 53–62.

Drew, A., and Ö. Karagedikli. 2008. "Some Benefits of Monetary Policy Transparency in New Zealand." Reserve Bank of New Zealand Discussion Paper Series No. 2008/01.

Ehrmann, M., and M. Fratzscher. 2004. "Taking Stock: Monetary Policy Transmission to Equity Markets." *Journal of Money, Credit, and Banking* 36 (4): 719–37.

———. 2007a. "Communication by Central Bank Committees—Different Strategies, Same Effectiveness?" *Journal of Money, Credit, and Banking* 39 (2–3): 509–41.

———. 2007b. "Transparency, Disclosure and the Federal Reserve." *International Journal of Central Banking* 3 (1): 179–225.

Geraats, P. M. 2005. "Transparency and Reputation: The Publication of Central Bank Forecasts." *B.E. Journal in Macroeconomics: Topics in Macroeconomics* 5 (1): 1–26.

Goodhart, C. A. E. 2001. "Monetary Transmission Lags and the Formulation of the Policy Decision on Interest Rates." *Review* (Federal Reserve Bank of St. Louis) 83 (4): 165–82.

Gravelle, T., and R. Moessner. 2002. "Reactions of Canadian Interest Rates to Macroeconomic Announcements: Implications for Monetary Policy Transparency." *Journal of Bond Trading and Management* 1: 27–43.

Gürkaynak, R. S., A. T. Levin, and E. T. Swanson. 2006. "Does Inflation Targeting Anchor Long-Run Inflation Expectations? Evidence from Long-Term Bond Yields in the US, UK and Sweden." Federal Reserve Bank of San Francisco Working Paper No. 2006–09.

Gürkaynak, R., B. Sack, and E. Swanson. 2005. "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements." *International Journal of Central Banking* 1 (1): 55–93.

Guthrie, G., and J. Wright. 2000. "Open Mouth Operations." *Journal of Monetary Economics* 46 (2): 489–516.

Ingves, S. 2007. "Communications — What Demands Are Made of an Independent Central Bank?" Speech at the Swedish Economics Association, Stockholm, April 24.

Issing, O. 2005. "Communication, Transparency, Accountability: Monetary Policy in the Twenty-First Century." *Review* (Federal Reserve Bank of St. Louis) 87 (2, part 1): 65–83.

King, M. 2007. "The MPC Ten Years On." Lecture delivered to the Society of Business Economists, London, May 2.

Kohn, D. 2005. "Central Bank Communications." Remarks at the annual meeting of the American Economic Association, Philadelphia, PA, January 9.

———. 2008. "Recent and Prospective Developments in Monetary Policy Transparency and Communications: A Global

Perspective." Remarks at the annual meeting of the Allied Social Science Associations, New Orleans, LA, January 5.

Kohn, D., and B. Sack. 2003. "Central Bank Talk: Does It Matter and Why?" Finance and Economics Discussion Series Working Paper No. 2003-55, Board of Governors of the Federal Reserve System.

Libich, J. 2006. "Inflexibility of Inflation Targeting Revisited: Modeling the 'Anchoring' Effect." CAMA Working Paper No. 2006-02, Australian National University.

McCaw, S., and S. Ranchhod. 2002. "The Reserve Bank's Forecasting Performance." *Reserve Bank Bulletin* 65 (4): 5–23.

Morris, S., and H. S. Shin. 2002. "Social Value of Public Information." *American Economic Review* 92 (5): 1521–34.

———. 2005. "Central Bank Transparency and the Signal Value of Prices." *Brookings Papers on Economic Activity* 2005 (2): 1–66.

Nelson, W. R., and R. Perli. 2005. "Selected Indicators of Financial Stability." *Irving Fisher Committee Bulletin* 23 (October): 92–105.

Ranchhod, S. 2003. "Comparison of Interest Rate Forecast Errors: Reserve Bank, NZIER and the National Bank of New Zealand." Mimeo, Reserve Bank of New Zealand.

Rosenberg, I. 2007. "Riksbank to Introduce Own Path for the Repo Rate." Speech at Danske Bank, Stockholm, January 17.

Rudebusch, G., and J. Williams. 2006. "Revealing the Secrets of the Temple: The Value of Publishing Interest Rate Projections." Mimeo, Board of Governors of the Federal Reserve Board.

Swanson, E. 2004. "Federal Reserve Transparency and Financial Market Forecasts of Short-Term Interest Rates." Finance and Economics Discussion Series Working Paper No. 2004-6, Board of Governors of the Federal Reserve System.

Svensson, L. E. O. 2004. "Challenges for Monetary Policy." Paper prepared for the Bellagio Group meeting at the National Bank of Belgium, Brussels, January 26–27.

———. 2006. "Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-transparency, Not Con." *American Economic Review* 96 (1): 448–52.

Tucker, P. 2007. "Central Banking and Political Economy: The Example of the UK's Monetary Policy Committee." Speech at the Inflation Targeting, Central Bank Independence and Transparency Conference, University of Cambridge, June 15.

Woodford, M. 2005. "Central-Bank Communication and Policy Effectiveness." Paper presented at the FRB Kansas City Symposium on "The Greenspan Era: Lessons for the Future," Jackson Hole, WY, August 25–27.