Real-Time Density Forecasts from VARs with Stochastic Volatility *

Todd E. Clark
Federal Reserve Bank of Kansas City

October 2009

Abstract

Central banks and other forecasters have become increasingly interested in various aspects of density forecasts. However, recent sharp changes in macroeconomic volatility — such as the Great Moderation and the more recent sharp rise in volatility associated with greater variation in energy prices and the deep global recession — pose significant challenges to density forecasting. Accordingly, this paper examines, with real-time data, density forecasts of U.S. GDP growth, unemployment, inflation, and the federal funds rate from VAR models with stochastic volatility. The model of interest extends the steady state prior BVAR of Villani (2009) to include stochastic volatility, because, as found in some prior work and this paper, incorporating informative priors on the steady states of the model variables often improves the accuracy of point forecasts. The evidence presented in the paper shows that adding stochastic volatility to a BVAR with some variables in gap form and a steady state prior materially improves the real-time accuracy of point and density forecasts.

JEL Nos.: C53, C32, E37
Keywords: Steady-state prior, Prediction, Bayesian methods

*Economic Research Dept.; Federal Reserve Bank of Kansas City; 1 Memorial Drive; Kansas City, MO 64198; `todd.e.clark@kc.frb.org`; phone: (816)881-2575; fax: (816)881-2199. The author gratefully acknowledges helpful conversations with Taeyoung Doh, John Geweke, Mattias Villani, and Chuck Whiteman and comments from Marek Jarocinski, Michael McCracken, Par Osterholm, Shaun Vahey, and seminar participants at the Federal Reserve Bank of Kansas City. Special thanks are due to Mattias Villani for providing simplified forms of GLS-based equations for the slope and steady state coefficients of the VAR with a steady state prior. The views expressed herein are solely those of the author and do not necessarily reflect the views of the Federal Reserve Bank of Kansas City or the Federal Reserve System.

1 Introduction

Policymakers and forecasters are increasingly interested in forecast metrics that require density forecasts of macroeconomic variables. Such metrics include confidence intervals, fan charts, and probabilities of recession or inflation exceeding or falling short of a certain threshold. For example, in 2008 the Federal Reserve expanded its publication of forecast information to include qualitative indications of the degree of uncertainty surrounding the outlook, including indications of whether uncertainty is higher or lower than usual, and more in one direction or the other. Other central banks, such as the Bank of Canada, Bank of England, Norges Bank, South African Reserve Bank, and Sveriges Riksbank, routinely publish fan charts that provide entire forecast distributions for inflation and, in some nations, a measure of output or the policy interest rate.

For many countries, however, changes in volatility over time pose a challenge to density forecasting. In the U.S., the Great Moderation significantly reduced the volatility of many macroeconomic variables. More recently, though, a variety of forces have substantially increased volatility. In the few years before the 2007-2009 recession, increased volatility of energy prices caused the volatility of total inflation to rise sharply. Then, the severe recession raised the volatility of a range of macroeconomic variables — by enough, as of this writing, to largely (although probably temporarily) reverse the Great Moderation in GDP growth.

Such shifts in volatility have the potential to result in forecast densities that are either far too wide or too narrow. For example, until recently the volatility of U.S. growth and inflation was much lower in data since the mid-1980s than in data for the 1970s and early 1980s. Consider, for example, forecasts of GDP growth. Density forecasts for, say, 2007, based on time series models assuming constant variances over a sample such as 1960-2006 would probably be far too wide. On the other hand, today, density forecasts for 2009 based on time series models assuming constant variances for 1985-2008 would probably be too narrow. Results in Jore, Mitchell, and Vahey (2009) support this intuition. In an analysis of real-time density forecasts since the mid-1980s, they find that models estimated with full samples of data and constant parameters fare poorly in density forecasting. Allowing discrete breaks in variances materially improves density forecasts made in the Great Moderation period.[1]

[1] In a similar vein, Hong, Li, and Zhao (2004) find that allowing time variation in volatilities (in their case through GARCH or Markov switching) significantly improves the accuracy of density forecasts of daily spot interest rates.

If volatility breaks were rare and always observed clearly with hindsight, simple split-sample or rolling sample methods might be used to obtain reliable density forecasts. But as recent events have highlighted, breaks such as the Great Moderation once thought to be effectively permanent can turn out to be shorter-lived, and reversed (at least temporarily). Over time, then, obtaining reliable density forecasts likely requires forecast methods that allow for routine breaks in volatilities.

Accordingly, this paper examines the accuracy of real-time density forecasts of macroeconomic variables made with vector autoregressions (VARs) that allow for continuous changes in the conditional variances of the model's shocks — that is, stochastic volatility. In light of the evidence in Clark and McCracken (2008, 2010) that the accuracy of point forecasts of U.S. GDP growth, inflation, and interest rates is improved by specifying the inflation and interest rates as deviations from a trend defined as a measure of long-run inflation expectations, the model of interest in this paper also specifies the unemployment rate, inflation, and interest rate variables in gap, or deviation from trend, form.[2]

In addition, based on a growing body of evidence on the accuracy of point forecasts, the VAR of interest incorporates an informative prior on the steady state values of the model variables. Villani (2009) develops a Bayesian estimator of a VAR with an informative prior on the steady state. Applications of the estimator in studies such as Adolfson, et al. (2007), Beechey and Osterholm (2008), and Osterholm (2008) have shown that the use of a prior on the steady state often improves the accuracy of point forecasts.[3] In a methodological sense, this paper extends the estimator of Villani (2009) to include stochastic volatility.[4]

Focusing on Bayesian VARs with stochastic volatility, most variables in gap form, and an informative prior on the steady state, this paper examines real-time point and density forecasts from a range of BVARs. The models include as variables (U.S.) GDP growth, unemployment, inflation, and the federal funds rate. A variety of approaches to modeling time variation in variances is considered: (1) assuming constant variances for the full sample, (2) assuming constant variances over a rolling 20 year sample used in model estimation, and (3) allowing stochastic volatility as in Cogley and Sargent (2005), Cogley, Morozov,

[2]The detrending used in these papers is motivated by the prior work of Kozicki and Tinsley (2001a,b) on moving endpoints.

[3]The relative entropy approach developed by Robertson, Tallman, and Whiteman (2005) could instead be used to impose desired steady states on forecasts.

[4]D'Agostino, Gambetti, and Giannone (2008) examine point forecasts from VARs with time-varying parameters and stochastic volatility.

and Sargent (2005), and Primiceri (2005).[5]

The evidence presented in the paper shows that adding stochastic volatility to the BVAR with most variables in gap form and a steady state prior materially improves the real-time accuracy of point and density forecasts. Compared to models with constant variances, models with stochastic volatility have lower RMSEs, significantly more accurate interval forecasts (coverage rates), probability integral transforms (PITs) that are closer to uniformity, normalized forecast errors (computed from the PITs) that are much closer to a standard normal distribution, and average log predictive density scores that are much lower.

Section 2 describes the real-time data used in the analysis. Section 3 presents the estimator of a VAR with stochastic volatility and an informative prior on the steady state means. Section 4 details the other forecasting models considered. Section 5 reports the results. Section 6 concludes.

2 Data

Forecasts are evaluated for four variables: output growth, the unemployment rate, inflation, and the federal funds rate. Output is measured as GDP or GNP (depending on data vintage). Inflation is measured with the GDP or GNP deflator or price index (depending on data vintage). Growth and inflation rates are measured as annualized log changes (from $t-1$ to t).

The raw quarterly data on output, prices, unemployment, and interest rates are taken from the Federal Reserve Bank of Philadelphia's Real-Time Data Set for Macroeconomists (RTDSM) and the Board of Governor's FAME database. Real-time data on GDP or GNP and the GDP or GNP price series are from the RTDSM. For simplicity, hereafter "GDP" and "GDP price index" refer to the output and price series, even though the measures are based on GNP and a fixed weight deflator for much of the sample. In the case of unemployment and fed funds rates, for which real-time revisions are small to essentially non–existent, I simply abstract from real-time aspects of the data. The quarterly data on unemployment and the interest rate are constructed as simple within-quarter averages of the source monthly data.

[5] Cogley, Morozov, and Sargent (2005) examine density forecasts for the U.K. from a VAR with time-varying parameters and stochastic volatility, but from a single point in time, as opposed to from the full historical perspective considered in this paper. Similarly, Beechey and Osterholm (2008) examine density forecasts for Australia from a model with an informative prior on the steady state, but also from a single point in time.

In the VARs with steady state priors, the unemployment rate, inflation, and funds rate variables are specified in gap, or deviation from trend, form, with the trends measured in real time. The trend specifications are based in part on the need to be able to easily and tractably account for the impact of trend uncertainty on the forecast distributions. Unemployment u_t is centered around a trend u^*_{t-1} computed by exponential smoothing, with a smoothing coefficient of 0.02: $u^*_t = u^*_{t-1} + 0.02(u_t - u^*_{t-1})$.

Inflation and the funds rate are centered around long-term inflation expectations from the Blue Chip Consensus — specifically, Consensus forecasts of average GDP price inflation 6-10 years ahead. The Blue Chip forecasts are taken from surveys published in the spring and fall of each year from 1979 through 2008. For model estimation purposes, the Blue Chip data are extended from 1979 back to 1960 with an estimate of expected GDP inflation based on exponential smoothing (with a smoothing parameter of .05). As noted by Kozicki and Tinsley (2001a,b) and Clark and McCracken (2008), exponential smoothing yields an estimate that matches up reasonably well with survey–based measures of long–run expectations in data since the early 1980s. The appendix provides additional detail on the real-time series of inflation expectations. Note that, to account for the uncertainty in the forecasts of inflation and the funds rate associated with the trend defined as the long-run inflation expectation, the VARs with steady state priors include the change in the expectation as an endogenous variable, which is forecast along with the other variables of the system.

The full forecast evaluation period runs from 1985:Q1 through 2008:Q3, which involves real-time data vintages from 1985:Q1 through 2009:Q1. As described in Croushore and Stark (2001), the vintages of the RTDSM are dated to reflect the information available around the middle of each quarter. Normally, in a given vintage t, the available NIPA data run through period $t-1$.[6] For each forecast origin t starting with 1985:Q1, I use the real-time data vintage t to estimate the forecast models and then construct forecasts for periods t and beyond. For forecasting models estimated recursively, the starting point of the model estimation sample is always 1961:Q1.

The results on forecast accuracy cover forecast horizons of 1 quarter ($h = 1Q$), 2 quarters ($h = 2Q$), 1 year ($h = 1Y$), and 2 years ($h = 2Y$) ahead. In light of the time $t-1$ information actually incorporated in the VARs used for forecasting at t, the 1-quarter ahead forecast forecast is a current quarter (t) forecast, while the 2-quarter ahead forecast is a next quarter

[6] In the case of the 1996:Q1 vintage, with which the BEA published a benchmark revision, the data run through 1995:Q3 instead of 1995:Q4. For this vintage, each of the forecast horizons are extended one period.

$(t+1)$ forecast. In keeping with Federal Reserve practice, the 1– and 2–year ahead forecasts for GDP growth and inflation are four–quarter rates of change (the 1–year ahead forecast is the percent change from period t through $t+3$; the 2–year ahead forecast is the percent change from period $t+4$ through $t+7$). The 1– and 2–year ahead forecasts for unemployment and the funds rate are quarterly levels in periods $t+3$ and $t+7$, respectively.

As discussed in such sources as Romer and Romer (2000), Sims (2002), and Croushore (2005), evaluating the accuracy of real-time forecasts requires a difficult decision on what to take as the actual data in calculating forecast errors. The GDP data available today for, say, 1985, represent the best available estimates of output in 1985. However, output as defined and measured today is quite different from output as defined and measured in 1970. For example, today we have available chain-weighted GDP; in the 1980s, output was measured with fixed-weight GNP. Forecasters in 1985 could not have foreseen such changes and the potential impact on measured output. Accordingly, I follow studies such as Romer and Romer (2000) and Faust and Wright (2007) and use the second available estimates of GDP/GNP and the GDP/GNP deflator as actuals in evaluating forecast accuracy. In the case of h–step ahead (for h = 1Q, 2Q, 1Y, and 2Y) forecasts made for period $t+h$ with vintage t data ending in period $t-1$, the second available estimate is normally taken from the vintage $t+h+2$ data set. In light of my abstraction from real-time revisions in unemployment and the funds rate, for these series the real-time data correspond to the final vintage data.

3 BVAR with stochastic volatility and informative priors on steady state means (BVAR-SSPSV)

The model of primary interest — henceforth denoted by BVAR-SSPSV (short for BVAR with most variables in gap form, an informative steady state prior, and stochastic volatility) — extends Villani's (2009) model with a steady state prior to include stochastic volatility, modeled as in Cogley and Sargent (2005).[7] The model can be estimated with a Metropolis-within-Gibbs Markov Chain Monte Carlo (MCMC) algorithm, combining (modified) por-

[7] While the importance of time variation in volatility compared to time variation in coefficients remains a matter of some debate (see, for example, Sims and Zha (2006) and the references therein), this model allows for some of both. To be sure, the modeling of volatility as stochastic could emphasize a role for time variation in volatility. However, the specification of most variables in gap or deviation from trend form allows the model to capture changes over time in the mean levels of unemployment, inflation, and the federal funds rate. The use of rolling sample estimates also allows, in a crude way, for some variation over time in coefficients.

tions of the algorithms of Villani (2009) and Cogley and Sargent (2005). This section details the model, estimation procedure, priors, and the generation of posterior distributions of forecasts from the model.

3.1 Model

Let y_t denote the $p \times 1$ vector of forecast variables and d_t denote a $q \times 1$ vector of deterministic variables. In this implementation, y_t includes GDP growth, the unemployment rate less its trend lagged one period, inflation less the long-run inflation expectation, the funds rate less the long-run inflation expectation, and the change in the long-run inflation expectation.[8] The only variable in d_t is a constant. Let $\Pi(L) = I_p - \Pi_1 L - \Pi_2 L^2 - \cdots - \Pi_k L^k$, Ψ = a $p \times q$ matrix of coefficients on the deterministic variables, and A = a lower triangular matrix with ones on the diagonal and coefficients a_{ij} in row i and column j (for $i = 2, \ldots, p$, $j = 1, \ldots, i-1$). The VAR(k) with stochastic volatility takes the form

$$\begin{aligned} \Pi(L)(y_t - \Psi d_t) &= u_t, \\ u_t &= A^{-1}\Lambda_t^{.5}\epsilon_t,\ \epsilon_t \sim N(0, I_p),\ \Lambda_t = \text{diag}(\lambda_{1,t}, \lambda_{2,t}, \lambda_{3,t}, \ldots, \lambda_{p,t}) \qquad (1) \\ \log(\lambda_{i,t}) &= \log(\lambda_{i,t-1}) + \nu_{i,t},\ \nu_{i,t} \sim \text{iid } N(0, \phi_i)\ \forall\ i = 1, \ldots, p. \end{aligned}$$

Under the stochastic volatility model, taken from Cogley and Sargent (2005), the log variances in Λ_t follow random walk processes. The (diagonal) variance-covariance matrix of the vector of innovations to the log variances is denoted Φ. This particular representation provides a simple approach to allowing time variation in the variances and covariances of the residuals u_t. While not important for forecasting purposes, it also allows recovery of structural shocks to monetary policy, under a recursive identification scheme. Under the above specification, the residual variance–covariance for period t is $\Sigma_t \equiv A^{-1}\Lambda_t A^{-1}$.

3.2 Estimation procedure

The model is estimated with a Metropolis-within-Gibbs MCMC algorithm, combining modified portions of the algorithms of Villani (2009) and Cogley and Sargent (2005). This subsection briefly describes the algorithm; details on priors are given in the next subsection.

Step 1: Draw the slope coefficients Π conditioned on Ψ, the history of Λ_t, A, and Φ.

[8] The inclusion of the long-run expectation as an endogenous variable does not appear to give the model with the steady state prior an advantage over the simple BVAR. A model without the expectation as an endogenous variable, in which the inflation expectation is assumed constant over the forecast horizon, yields results similar to those reported for the BVAR-SSP specifications.

Recast the demeaned VAR in state–space form, where the measurement vector Y_t is defined as $y_t - \Psi d_t$, X_t contains the appropriate lags of $y_t - \Psi d_t$, and vec(Π) is the vector of VAR slope coefficients:

$$Y_t = (I_p \otimes X_t') \cdot \text{vec}(\Pi) + u_t,\ var(u_t) = \Sigma_t = A_t^{-1}\Lambda_t A_t^{-1}. \tag{2}$$

In a GLS extension of Villani (2009), the vector of coefficients is sampled from a normal posterior distribution with mean $\bar{\mu}_\Pi$ and variance $\bar{\Omega}_\Pi$, based on prior mean μ_Π and Ω_Π, where:[9]

$$\bar{\Omega}_\Pi^{-1} = \Omega_\Pi^{-1} + \sum_{t=1}^{T}(\Sigma_t^{-1} \otimes X_t X_t') \tag{3}$$

$$\bar{\mu}_\Pi = \bar{\Omega}_\Pi \left\{ \text{vec}(\sum_{t=1}^{T} \Sigma_t^{-1} Y_t X_t') + \Omega_\Pi^{-1}\mu_\Pi \right\}. \tag{4}$$

Step 2: Draw the steady state coefficients Ψ conditioned on Π, the history of Λ_t, A, and Φ.

For this step, the VAR is rewritten as

$$q_t = \Pi(L)\Psi d_t + u_t,\ \text{where } q_t \equiv \Pi(L) y_t. \tag{5}$$

The dependent variable q_t is obtained by applying to the vector y_t the lag polynomial estimated with the preceding draw of the Π coefficients. The right-hand side term $\Pi(L)\Psi d_t$ simplifies to $\Theta \bar{d}_t$, where, as in Villani (2009) with some modifications, $\bar{d}_t$ contains current and lagged values of the elements of d_t, and Θ is defined such that vec(Θ) = Uvec(Ψ). With four lags in the VAR, $\bar{d}_t$ and U are

$$\bar{d}_t = (d_t', -d_{t-1}', -d_{t-2}', -d_{t-3}', \ldots, -d_{t-k}')' \tag{6}$$

$$U = \begin{pmatrix} I_{pq \times pq} \\ I_q \otimes \Pi_1 \\ I_q \otimes \Pi_2 \\ I_q \otimes \Pi_3 \\ \vdots \\ I_q \otimes \Pi_p \end{pmatrix}. \tag{7}$$

In a GLS extension of Villani (2009), the vector of coefficients Ψ is sampled from a normal posterior distribution with mean $\bar{\mu}_\Psi$ and variance $\bar{\Omega}_\Psi$, based on prior mean μ_Ψ and

[9] Special thanks are due to Mattias Villani for providing the formulae for posterior means and variances of Π and Ψ, which greatly simplified an approach I used when starting this paper.

Ω_Ψ, where:

$$\bar{\Omega}_\Psi^{-1} = \Omega_\Psi^{-1} + U'\Big\{\sum_{t=1}^{T}(\bar{d}_t\bar{d}_t' \otimes \Sigma_t^{-1})\Big\}U \tag{8}$$

$$\bar{\mu}_\Psi = \bar{\Omega}_\Psi\left\{U'\text{vec}(\sum_{t=1}^{T}\Sigma_t^{-1}q_t\bar{d}_t') + \Omega_\Psi^{-1}\mu_\Psi\right\}. \tag{9}$$

Step 3: Draw the elements of A conditioned on Π, Ψ, the history of Λ_t, and Φ.

Following Cogley and Sargent (2005), rewrite the VAR as

$$A\Pi(L)(y_t - \Psi d_t) \equiv A\hat{y}_t = \Lambda_t^{.5}\epsilon_t, \tag{10}$$

where, conditioned on Π and Ψ, $\hat{y}_t$ is observable. This system simplifies to a set of $i = 2, \ldots, p$ equations, with equation i having as dependent variable $\hat{y}_{i,t}$ and as independent variables $-1 \cdot \hat{y}_{j,t}, j = 1, \ldots, i-1$, with coefficients a_{ij}. Multiplying each equation i through by $\lambda_{i,t}^{-.5}$ eliminates the heteroskedasticity associated with stochastic volatility. Then, proceeding separately for each transformed equation i, draw the i'th equation's vector of j coefficients a_{ij} from a normal posterior distribution with the mean and variance implied by the posterior mean and variance computed in the usual (OLS) way. See Cogley and Sargent (2005) for details.

Step 4: Draw the elements of the variance matrix Λ_t conditioned on Π, Ψ, A, and Φ.

Following Cogley and Sargent (2005), the VAR can be rewritten as

$$A\Pi(L)(y_t - \Psi d_t) \equiv \tilde{y}_t = \Lambda_t^{.5}\epsilon_t, \tag{11}$$

where $\epsilon_t \sim N(0, I_p)$. Taking logs of the squares yields

$$\log \tilde{y}_{i,t}^2 = \log \lambda_{t,t} + \log \epsilon_{i,t}^2, \ \forall\ i = 1, \ldots, p. \tag{12}$$

The conditional volatility process is

$$\log(\lambda_{i,t}) = \log(\lambda_{i,t-1}) + \nu_{i,t}, \ \nu_{i,t} \sim \text{iid } N(0, \phi_i)\ \forall\ i = 1, \ldots, p. \tag{13}$$

The estimation of the time series of $\lambda_{i,t}$ proceeds equation by equation, using the measured $\log \tilde{y}_t^2$ and Cogley and Sargent's (2005) version of the Metropolis algorithm of Jacquier, Polson, and Rossi (1994). See Cogley and Sargent (2005) for details.

Step 5: Draw the innovation variance matrix Φ conditioned on Π, Ψ, the history of Λ_t, and A.

Following Cogley and Sargent (2005), the sampling of the diagonal elements of Φ, the variances of innovations to log volatilities, is based on inverse Wishart priors and posteriors. For each equation i, the posterior scaling matrix is a linear combination of the prior and the sample variance innovations computed as the variance of $\tilde{\lambda}_{i,t} - \tilde{\lambda}_{i,t-1}$. I obtain draws of each Φ_i by sampling from the inverse Wishart posterior with this scale matrix.

3.3 Priors and other estimation details

The prior for the VAR slope coefficients $\Pi(L)$ is based on a Minnesota specification. The prior means suppose each variable follows an AR(1) process (so all prior means except those on the own first lag are 0), with coefficients of 0.25 for GDP growth and 0.8 for the other variables.[10] Prior standard deviations are controlled by the usual hyperparameters, with overall tightness of 0.2, cross–equation tightness of 0.5, and linear decay in the lags. The standard errors used in setting the prior are obtained from training sample estimates of univariate AR(4) models, where the training sample consists of the 40 observations preceding the estimation sample used for a given vintage.[11]

Priors are imposed on the deterministic coefficients Ψ to push the long-run means and forecasts of the model variables toward certain values. Specifically, steady state priors are set to push the means/long-run forecasts toward: (1) GDP growth, 3.0 percent; (2) unemployment less the exponentially smoothed trend, 0.0; (3) inflation less the long-run inflation expectation of Blue Chip, 0.0; (4) federal funds rate less the long-run inflation expectation of Blue Chip, 2.5; and (5) change in the long-run inflation expectation of Blue Chip, 0.0. Accordingly, in the prior for the elements of Ψ, all means are zero, except as follows: GDP growth, intercept coefficient of 3.0; and fed funds rate, intercept coefficient of 2.5. In the recursive (rolling) estimation, I set the following standard deviations on the steady-state means: GDP growth, 0.2 (0.3); unemployment less trend, 0.2 (0.4); inflation less long-run expectation, 0.2 (0.3); fed funds rate less long-run inflation expectation, 0.6

[10] The prior mean on the AR(1) coefficient for GDP growth pushes the VAR equation toward an AR(1) process that reasonably characterizes U.S. data. The prior means on the other variables push the equations toward processes that are significantly more persistent, but short of unit root persistence. The specification of the unemployment rate, inflation, and the federal funds rate in deviation from trend form generally reduces persistence.

[11] Because the federal funds rate data do not begin until mid-1954 and the market was thin in early years, the interest rate series for 1950:Q1-1959:Q4 (the earliest training sample regression period is 1951:Q1 through 1960:Q4) is defined as the 3-month Treasury bill rate. For those data vintages in which output and price index data are not available for the pre-sample period, the pre-sample estimates are simply taken to be those from the most recent (earlier) vintage with data available for the pre-sample period.

(0.75); and change in long-run inflation expectation, 0.2 (0.2).[12]

As to the priors on the stochastic volatility portion of the model, I use uninformative priors for the elements of A and loose priors for the initial values of $\log(\lambda_{i,t})$ and the variances of the innovations to $\log(\lambda_{i,t})$. More specifically, I use the following priors:

$$\begin{aligned} \log \lambda_{i,0} &\sim N(\log \hat{\lambda}_{i,OLS},\ 4)\ \forall\ i = 1, \ldots, p \\ a_i &\sim N(0, 1000^2 \cdot I_{i-1})\ \forall\ i = 2, \ldots, p \\ \phi_i &\sim IW(5 \cdot 0.035,\ 5)\ \forall\ i = 1, \ldots, p, \end{aligned}$$

where a_i denotes the $(i-1) \times 1$ vector of $a_{i,j}$ coefficients in the i'th row of A and the $\hat{\lambda}_{i,OLS}$ are obtained from the residuals of AR(4) models estimated with a training sample of the 40 observations preceding the estimation sample. The variance of 4 on each $\log \lambda_{i,0}$ corresponds to a quite loose prior on the initial variances, in light of the log transformation of the variances.

3.4 Drawing forecasts

For each (retained) draw in the MCMC chain, I draw forecasts from the posterior distribution using an approach like that of Cogley, Morozov, and Sargent (2005). To incorporate uncertainty associated with time variation in Λ_t over the forecast horizon of 8 periods, I sample innovations to Λ from a normal distribution with (diagonal) variance Φ, and use the random walk specification to compute Λ_t from Λ_{t-1}. For each period of the forecast horizon, I then sample shocks to the VAR with a variance of Σ_{t+h} and compute the forecast draw of Y_{t+h} from the VAR structure and drawn shocks.

In all forecasts obtained from models with steady state priors, the model specification readily permits the construction of forecast distributions that account for the uncertainty associated with the trend unemployment rate and long-run inflation expectation. (Recall that, as described in section 2 and detailed in the appendix, at each forecast origin, the inflation expectation series is measured with the vintage of data available in real time.) In each draw, the model is used to forecast GDP growth, unemployment less trend lagged one period, inflation less the long-run inflation expectation, the funds rate less the long-run inflation expectation, and the change in the long-run inflation expectation. The forecasted changes in the long-run expectation are accumulated and added to the value at the end of

[12] For most of the variables, I use slightly tighter priors for the recursive scheme than the rolling because, in the recursive case, the gradual increase in the size of the estimation sample (as forecasting moves forward) gradually reduces the influence of the prior.

the estimation sample to obtain the forecasted level of the expectation. The forecasts of the level of the expectation are then added to the forecasts of inflation less the expectation and the funds rate less the expectation to obtain forecasts of the levels of inflation and the funds rate. Forecasts of the level of the unemployment rate and the exponentially smoothed trend are obtained by iterating forward, adding the lagged trend value to obtain the forecast of the unemployment rate, then computing the current value of the unemployment trend, and continuing forward in time over the forecast horizon.

Finally, I report posterior estimates based on a sample of 10,000 draws, obtained by first generating 10,000 burn-in draws and then saving every fifth draw from another 50,000 draws.[13] Point forecasts are constructed as posterior means of the MCMC distributions.

4 Other Models Considered

To establish the effectiveness of steady state priors and stochastic volatility, forecasts from the BVAR-SSPSV model are compared against a range of forecasts from other models. In light of the evidence in Clark and McCracken (2008, 2010) that point forecasts from VARs are often dominated by point forecasts from univariate models, the set of models considered in the evaluation of point (not density) forecasts includes univariate specifications as benchmarks. The set of models also includes conventional BVARs without steady state priors or stochastic volatility and BVARs with steady priors and not stochastic volatility. This section provides details on these other models.

4.1 Univariate models

For output, widely modeled as following low-order AR processes, the univariate model is an AR(2). The univariate model for unemployment is an AR(2) in the change in the unemployment rate. In the case of inflation, I follow Stock and Watson (2007) and use an MA(1) process for the change in inflation, estimated with a rolling window of 40 observations. Stock and Watson find that the IMA(1) generally outperforms random walk or AR model forecasts of inflation. In light of some general similarities in the time series properties of inflation and short–term interest rates and the IMA(1) rationale for inflation described by Stock and Watson, the univariate model for the short-term interest rate is also specified as

[13] For a model with stochastic volatility and time-varying parameters, Cogley and Sargent (2005) only save every tenth draw to reduce the correlation in the MCMC chain. In this analysis, with time-invariant VAR coefficients, I save only every fifth draw, to speed up the model estimation that is done for nearly 100 vintages of data (i.e., nearly 100 times).

an MA(1) in the first difference of the series, estimated with a rolling window of 10 years of data (the IMA(1) generally outperforms AR model forecasts of the funds rate).

4.2 Simple BVARs

One multivariate forecasting model is a BVAR(4), in GDP growth, the unemployment rate, inflation, and the federal funds rate. The model is estimated with Minnesota priors — specifically the Normal–diffuse prior described in Kadiyala and Karlsson (1997). The prior means and variances (determined by hyperparameters) are the same as described in section 3.3 for the BVAR-SSPSV model. Flat priors are used for the intercepts of the equations. The model is estimated with Gibbs sampling, as described in Kadiyala and Karlsson (1997). I consider both full sample and rolling sample estimates of the model and forecasts, using a rolling sample of the most recent 20 years of data. The number of draws is 15,000, with the first 5000 discarded. The rolling sample serves as a crude approach to capturing the potential impacts of changing volatility on appropriate forecast confidence intervals.

4.3 BVARs with steady state prior (BVAR-SSP)

I also consider forecasts from a BVAR(4) with most variables in gap form and an informative prior on the steady state, as specified and estimated in Villani (2009). The model variables consistent of GDP growth, the unemployment rate less its trend lagged one period, inflation less the long-run inflation expectation, the funds rate less the long-run inflation expectation, and the change in the long-run inflation expectation. Using the notation above, the model takes the form

$$\Pi(L)(y_t - \Psi d_t) = u_t,\ u_t \sim N(0, \Sigma), \tag{14}$$

with four lags. Using the Minnesota and steady state priors described above for the model also including stochastic volatility (along with a diffuse prior on Σ), I estimate the model with the Gibbs sampling approach given in Villani (2009). The estimates and forecasts are obtained from a total of 15,000 draws, with the first 5000 discarded. I consider forecasts from the model estimated under two different schemes, one using the full sample of data available at the forecast origin (recursive), and the other using a rolling sample of the most recent 20 years of data (rolling).

5 Results

For the models with stochastic volatility to yield density forecasts more accurate than those from models with constant volatilities, it likely needs to be the case that volatility has varied significantly over time. Therefore, as a starting point, it is worth considering the estimates of stochastic volatilities from the BVAR-SSPSV model — specifically, time series of residual standard deviations (posterior means of the time series of standard deviations) estimated under the recursive scheme. To first establish the properties of the estimator over time in the absence of source data revisions, Figure 1 reports estimates based on just the last available vintage of data, the 2009:Q1 vintage. For each variable (except that, for brevity, the chart focuses on the primary variables of interest and omits the volatility estimates for the change in the long-run inflation expectation), the red line provides the volatility time series estimated with data from 1961 through 2008. The green, blue, and black lines provide time series estimated with data samples ending in, respectively, 1998:Q4, 1991:Q4, and 1984:Q4 (using pseudo-vintages dated 1999:Q1, 1992:Q1, and 1985:Q1). Overall, the estimates confirm significant time variation in volatility, and generally match the contours of estimates shown in such studies as Cogley and Sargent (2005). In particular, volatility fell sharply in the mid-1980s with the Great Moderation. The estimates also reveal a sharp rise in volatility in recent years, reflecting the rise in energy price volatility and the sharp recession of 2007-2008. In fact, in the case of inflation, volatility has risen to pre-Moderation levels; for GDP growth, volatility has retraced much, although not all, of the Great Moderation.

Across estimates generated with different data samples, the volatilities are generally very similar (that is, the different colors of lines are very similar). Not surprisingly, the biggest revisions occur with volatilities estimated at the end of a sample (the biggest differences in lines occur at the end of the lines) — only being able to do one-sided filtering at the end of a sample has some modest effect. For example, in the case of GDP growth, for volatility in 1990, the estimate obtained with a data sample ending in 1991 (blue line) exceeds the estimate obtained with a data sample ending in 2008.

As might be expected, comparing estimates across real-time data vintages yields larger changes in volatility estimates across vintages. Figure 2 shows corresponding time series of volatility estimates, but obtained with real-time data vintages instead of just the final data vintage. Data revisions — driven by benchmark revisions and large annual revisions

— lead to more significant differences across samples/vintages in the stochastic volatility estimates for GDP growth and GDP inflation, but with little impact for unemployment and the funds rate. For growth and inflation, the general contours of volatility are very similar across vintages, but levels can differ somewhat. It remains to be seen whether such changes in real time estimates are so great as to make it difficult to incorporate stochastic volatility and still improve the accuracy of point and density forecasts.

This section proceeds first with results for real-time point forecasts, in the form of mean errors and RMSEs. The following subsections presents results for density forecasts: probabilities of forecasts falling within 70 percent confidence intervals, histograms of probability integral transforms (PITs), the tests of Berkowitz (2001) applied to normal transforms of the PITs, and log predictive scores. To further illustrate the practical consequences of time variation in conditional volatilities, the section concludes by presenting fan charts of forecasts made in 2008:Q3. Recall that, as detailed in section 3.4, the forecast distributions for the BVAR-SSP and BVAR-SSPSV models account for the uncertainty associated with the unemployment trend and long-run inflation expectation (through the exponential smoothing model of the unemployment trend and the inclusion of the long-run inflation expectation as an endogenous variable).

5.1 Point forecasts

Tables 1 and 2 present real-time mean forecast errors and RMSEs for 1985-2008:Q3 (even though the BVAR-SSP and BVAR-SSPSV models include the long-run expectation as an endogenous variable, the presented results exclude this variable, and focus on the four variables of primary interest). In the case of the RMSEs, the first block reports RMSEs for univariate model forecasts; the remaining blocks report ratios of RMSEs for a given forecast model or method relative to the univariate model for a given variable. In these blocks, entries with value less than 1 mean a VAR forecast is more accurate than the univariate benchmark. Note that, for the VARs, point forecasts are defined as the means of the posterior distributions obtained by MCMC.

The mean errors of forecasts for growth, inflation, and the funds rate are consistently negative, indicating the forecasts are consistently too high. On this dimension, the rolling VARs often fare better than the recursively estimated VARs, with smaller (in absolute value) mean errors. For example, in the case of 1-year ahead forecasts of GDP growth, the mean error from the recursive BVAR-SSP is -0.540, while the mean error from the rolling BVAR

is -0.064. In general, the average errors in forecasts of unemployment are materially smaller than the average errors for other variables, but are more mixed in sign (sometimes positive, sometimes negative). For instance, with the recursive BVAR-SSP model, the 1-year ahead forecast errors average -0.540 for GDP growth and 0.049 for unemployment.

In terms of average errors at longer horizons, the BVARs with most variables in gap form and informative steady state priors (for simplicity, much of the discussion below simply refers to these models as BVARs with steady state priors) tend to fare better than the conventional BVARs, especially for inflation and the funds rate. For 2-year ahead projections of inflation, the rolling BVAR yields an average error of -1.093, compared to the rolling BVAR-SSP's mean error of -0.446. The BVARs with steady state priors and stochastic volatility tend to yield average errors broadly comparable to the BVARs with just steady state priors — generally a bit lower in the recursive case but sometimes a bit higher in the rolling case. In the case of 1-quarter ahead forecasts of GDP growth, the recursive (rolling) BVAR-SSP has an average error of -0.703 (-0.192), while the recursive (rolling) BVAR-SSPSV has an average error of -0.389 (-0.249). Finally, in many but not all cases, the mean errors are smaller (in absolute value) for the univariate forecasts than the VAR forecasts.

Consistent with the findings of Clark and McCracken (2008, 2010), by the metric of RMSEs the performance of the conventional BVARs (without variables in gap form and without steady state priors) relative to the univariate models is mixed. For example, at horizons of 1 and 2 quarters and 1 year, the BVAR forecasts often have RMSEs in excess of the univariate RMSE. But for growth, inflation, and the funds rate, the accuracy of BVAR forecasts relative to the univariate forecasts improves as the forecast horizon increases. At the 2-year horizon, BVAR forecasts of these variables are almost always more accurate than univariate forecasts, often substantially. Consider forecasts of unemployment from the recursive BVAR: the RMSE ratio declines from 1.047 at the 1-quarter horizon to 0.989 at the 1-year horizon to 0.722 at the 2-year horizon.

While the pattern is not entirely uniform, for the most part BVARs estimated with rolling samples yield lower RMSEs than BVARs estimated recursively (the pattern is clearer in the set of models with steady state priors). As examples, the RMSE ratios of 1-year ahead forecasts of GDP growth are 1.059 with the recursive BVAR-SSP and 0.954 with the rolling BVAR-SSP, and the RMSE ratios of 1-year ahead forecasts of unemployment are 0.949 and 0.877 with, respectively, the recursive and rolling BVAR-SSP specifications.

The BVARs with most variables in gap form and steady state priors generally yield lower RMSEs than conventional BVARs.[14] The advantage is probably most striking for 2-year ahead forecasts of inflation. Under a rolling estimation scheme, BVAR and BVAR-SSP forecasts have RMSE ratios of 1.805 and 1.078, respectively. But the advantage, albeit smaller, also applies in most other cases — for other variables and shorter horizons. At the 1-quarter horizon, rolling BVAR and BVAR-SSP forecasts of GDP growth have RMSE ratios of 1.148 and 1.091, respectively. At the 2-quarter horizon, rolling BVAR and BVAR-SSP forecasts of the funds rate have RMSE ratios of 1.083 and 1.006, respectively.

Adding stochastic volatility to the BVARs with most variables in gap form and steady state priors tends to further improve forecast RMSEs.[15] Such a finding appears to be consistent with Sims and Zha's (2006) emphasis on the importance of variance shifts and potential impacts on coefficient estimates. At the 1-quarter horizon, the recursive BVAR-SSP yields RMSE ratios of 1.157 for GDP growth and 1.203 for the funds rate, while the recursive BVAR-SSPSV yields corresponding ratios of 1.074 and 1.008. At the 1-year horizon, the recursive BVAR-SSP yields RMSE ratios of 1.059 for GDP growth and 0.997 for the funds rate, while the recursive BVAR-SSPSV yields corresponding ratios of 0.984 and 0.914. By the RMSE metric, the rolling BVAR-SSPSV is probably the single best model, beating the accuracy of univariate forecasts at all horizons for all variables, with the exception of GDP growth at the 1 and 2 quarter horizons, and effectively equaling or beating the accuracy of the other VAR models. The recursive BVAR-SSPSV is comparable, but on balance probably not quite as accurate as the rolling version of the specification.

5.2 Density forecasts: interval forecasts

In light of central bank interest in uncertainty surrounding forecasts, confidence intervals, and fan charts, interval forecasts — that is, coverage rates — provide a natural starting point for forecast density evaluation. Recent studies such as Giordani and Villani (2008) have used interval forecasts as a measure of forecast accuracy for macroeconomic density forecasts. Table 3 reports the frequency with which actual real-time outcomes for growth, unemployment, inflation, and the funds rate fall inside 70 percent highest posterior density intervals estimated in real time with the BVARs. Accurate intervals should result in fre-

[14] This finding is in line with prior evidence in Clark and McCracken (2008, 2010) on the advantage of detrending and evidence in Adolfson, et al. (2007), Beechey and Osterholm (2008), and Osterholm (2008) on the advantage of steady state priors.

[15] D'Agostino, Gambetti, and Giannone (2008) find that allowing time-varying parameters and stochastic volatility improves the accuracy of point forecasts from VARs.

quencies of about 70 percent. A frequency of more (less) than 70 percent means that, on average over a given sample, the posterior density is too wide (narrow). The table includes p-values for the null of correct coverage (empirical = nominal rate of 70 percent), based on t-statistics (using OLS variance estimates at the 1-step horizon and Newey-West variances with bandwidth $1.5 \cdot$ horizon in other cases).[16] These p-values are provided as a rough gauge of the importance of deviations from correct coverage. Note that one reason the p-values provide only a rough gauge is that the theory underlying Christofferson's (1998) test results abstracts from forecast model estimation — that is, parameter estimation error — while all forecasts considered in this paper are obtained from estimated models.

As Table 3 shows, the BVAR and BVAR-SSP intervals tend to be too wide, with actual outcomes falling inside the intervals much more frequently than the nominal 70 percent rate. For example, for the 1-quarter ahead forecast horizon, the recursive BVAR-SSP coverage rates (for actuals falling inside the 70 percent posterior interval) range from 84.2 to 94.7 percent. Based on the reported p-values, all of these departures from the nominal coverage rate appear to be statistically meaningful. Using the rolling estimation scheme yields slightly to somewhat more accurate interval forecasts (but the departures remain large enough to deliver low p-values, with the exception of the inflation forecasts), with BVAR-SSP coverage rates ranging from 73.7 to 90.5 percent at the 1-step ahead horizon. For most variables, the interval forecasts become more accurate at the 1-year horizon, with coverage rates closer to 70 percent; the inflation forecasts are the exception. For example, in the case of unemployment forecasts from the rolling BVAR-SSP, the coverage rate improves from 84.2 percent at the 1 quarter horizon to 78.3 at the 1 year horizon; in the case of inflation forecasts, the coverage rate deteriorates from 73.7 at the 1 quarter horizon to 80.4 at the 1 year horizon.

Adding stochastic volatility to the BVAR with a steady state prior materially improves the accuracy of interval forecasts. For the 1-quarter ahead forecast horizon, the rolling BVAR-SSPSV coverage rates range from 69.5 to 78.9 percent, compared to the rolling BVAR-SSP's range of 73.7 to 90.5 percent. With the stochastic volatility specifications, for growth, unemployment, and inflation forecasts the p-values for 1-step ahead coverage all exceed 40 percent (90 percent in all but one case). But coverage remains too high in the case of the funds rate, at roughly 80 percent — materially better than in the models

[16] At the one-step horizon, p-values from Christofferson's (1998) likelihood ratio test for unconditional coverage are very similar to those reported.

without stochastic volatility, but still too high. At the 1-year ahead horizon, the rolling BVAR-SSPSV coverage rates range from 69.6 to 77.2 percent, compared to the rolling BVAR-SSP's range of 78.3 to 83.7 percent.

5.3 Density forecasts: PITs

The probability integral transform (PIT) emphasized by Diebold, Tay, and Gunther (1998) provides a more general indicator of the accuracy of density intervals. For each variable and the forecast horizon of 1 quarter, Figures 3-6 present PIT histograms, obtained as decile counts of PIT transforms. For optimal density forecasts at the 1-step horizon, the PIT series would be independent uniform (0,1) random variables. Accordingly, the histograms would be flat (with 9.5 observations per bin at the 1-quarter horizon). Studies such as Christoffersen and Mazzotta (2005), Clements (2004), Geweke and Amisano (2008), and Hong, Li, and Zhao (2004) consider similar measures of density forecasts. To provide some measure of a gauge of the importance of departures from the iid uniform distribution, Figures 3-6 also include 95 percent intervals estimated under the binomial distribution (following Diebold, Tay, and Gunther (1998)). Again, these intervals are only intended as a rough guide; among other issues, the intervals abstract from the possible effects of model parameter estimation on the large-sample distributions of PITs.

Forecasts from BVARs without stochastic volatility suffer seemingly large departures from uniformity. While some might see the PITs for rolling BVARs as looking a bit flatter than those from recursive BVARs, the differences are pretty small, with both suffering material departures from uniformity. In the case of 1-quarter ahead GDP growth and unemployment rate forecasts, the PITs have too much mass in the middle of the distribution. The PITs for inflation are somewhat flatter, with more modest crossings of the 95 percent bands. The departures from uniformity are most severe for the PITs of funds rate forecasts; the histograms of the BVAR-SSP forecast PITs look more like normal densities than uniform densities. The clustering of mass in the middle of the distributions of the PITs most likely reflects estimated forecast distributions that are too wide, because the forecast models without stochastic volatility treat the residual variances as constant and therefore, following the Great Moderation, over-estimate volatilities.

Forecasts from BVARs with stochastic volatility look to be much closer to being uniformly distributed. For the recursive and rolling BVAR-SSPSV forecasts, the PITs are quite a bit flatter than for the models without stochastic volatility; the BVAR-SSPSV PITs

are much less prone to crossing the 95 percent bands. The BVAR-SSPSV PITs exceed the threshold at most once each for unemployment and the funds rate, and never for GDP growth and inflation.

5.4 Density forecasts: normal transforms of PITs

Normal transforms of the PITs can also provide useful indicators of the accuracy of density forecasts. The normalized forecast error is defined as $\Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of a one-step ahead forecast error and Φ^{-1} is the inverse of the standard normal distribution function. As developed in Berkowitz (2001), the normalized forecast error should be an independent standard normal random variable, because the PIT series should be an independent uniform(0,1) random variable. Berkowitz develops tests based on the normality of the normalized errors that have better power than tests based on the uniformity of the PITs. These tests have been used in recent studies such as Clements (2004) and Jore, Mitchell, and Vahey (2009).[17] Giordani and Villani (2008) also suggest that time series plots of the normalized forecast errors provide useful qualitative evidence of forecast density accuracy, and may reveal advantages or disadvantages of a forecast not evident from alternatives such as PIT histograms.

Figures 7-10 report time series of the normalized forecast errors, with bands representing 95 percent intervals for the normal distribution. In line with the PITs results, normalized errors from BVARs without stochastic volatility suffer seemingly important departures from the standard normal distribution. Many of the charts indicate the normalized errors have variances well below 1, non-zero means, and serial correlation. The most dramatic examples are for forecasts of the funds rate — such as from the recursive BVAR-SSP. Less dramatic, although still clear, examples include forecasts of GDP growth and unemployment from the recursive BVAR-SSP. The transforms look best (closest to the standard normal conditions) for forecasts of GDP inflation, which are clearly more variable. In some cases, results seem to look a bit better — at least in the sense of having a larger variance — under the rolling estimation scheme than the recursive, but qualitatively similar.

The normalized forecast errors from BVARs with stochastic volatility look much better — with larger variances and means closer to zero. In the case of GDP growth, variability of normalized errors is clearly greater for the BVAR-SSPSV specifications than the BVAR-

[17] Here, too, though, it is not known how model parameter estimation may affect the properties of such tests.

SSP or BVAR models, and the mean also looks to be closer to zero.[18] However, even with stochastic volatility, there remains an extended period of negative errors in the early 1990s, which implies serial correlation in the errors. The same basic pattern applies to the normalized errors of unemployment forecasts. The results in Figure 9 for inflation forecasts also suggest stochastic volatility improves the behavior of normalized errors, although not as dramatically as with GDP growth and unemployment. Finally, in the case of funds rate forecasts, allowing stochastic volatility also significantly increases the variance of the normalized errors, but seems to leave strong serial correlation.

To more formally document what the plots suggest, Table 4 reports various test metrics: the variances of the normalized errors, along with p-values for the null that the variance equals 1; the means of the normalized errors, along with p-values for the null of a zero mean; the AR(1) coefficient estimate and its p-value, obtained by a least squares regression including a constant; and the p-value of Berkowitz's (2001) likelihood ratio test for the joint null of a zero mean, unity variance, and no (AR(1)) serial correlation.

The tests confirm that, without stochastic volatility, variances are materially below 1, means are sometimes non-zero, and serial correlation can be considerable. For example, with the recursive BVAR-SSP model, the variances of the normalized forecast errors range from 0.206 (funds rate) to 0.637 (inflation), with p-values close to 0. With the same model, the AR(1) coefficients are 0.310 for GDP growth, 0.395 for unemployment, -0.199 for inflation, and 0.673 for the funds rate; the corresponding p-values are all close to zero, except in the case of inflation, for which the p-value is 0.060. However, particularly in terms of means and variances, the rolling scheme fares somewhat better than the recursive. Not surprisingly, given results such as these for means, variances, and AR(1) coefficients, the p-values of the Berkowitz (2001) test are nearly zero for all recursive and rolling BVAR-SSP forecasts, with the exception of rolling forecasts of GDP inflation.

By the formal metrics, as by the charts, allowing stochastic volatility improves the accuracy of density forecasts. In the case of the recursive BVAR-SSPSV specification, the variances of the normalized forecast errors range from 0.848 (federal funds rate) to 1.030 (inflation), with p-values of 0.223 or more. The AR(1) coefficients are all lower (in absolute value) for forecasts from the recursive BVAR-SSPSV than from the recursive BVAR-SSP

[18] Qualitatively, Giordani and Villani (2008) obtain quite similar results in comparing forecasts of GDP growth from a constant parameter AR model to forecasts from a model that allows coefficient and variance breaks.

specification. For unemployment and inflation (not growth or the funds rate), the p-values of the Berkowitz (2001) test are above 10 percent. For GDP growth, the p-value of the test exceeds 5 percent.

5.5 Density forecasts: log predictive density scores

The overall accuracy of the density forecasts can most broadly measured with log predictive density scores, used in such recent studies as Adolfson, Linde, and Villani (2005) and Geweke and Amisano (2008). For computational tractability, I compute the log predictive density score based on the Gaussian (quadratic) formula given in Adolfson, Linde, and Villani (2005), under which a lower score implies a better model. In addition to reporting average log scores for the full vector of variables, I report scores for each individual variable.

To help provide a rough gauge of the significance of score differences, I rely on the methodology developed in Amisano and Giacomini (2007), and report p-values for selected differences in mean scores, under the null of a zero mean. The variance of the mean is computed with a Newey-West HAC estimator, using a bandwidth of 2 for 1-quarter ahead forecasts and 1.5*horizon for other forecasts. Because the theoretical basis for the test provided by Amisano and Giacomini requires forecasts estimated with rolling samples of data (but does take account of the effects of uncertainty associated with the estimation of model parameters), I only apply the test to pairs of forecasts from models estimated with the rolling scheme: BVAR against BVAR-SSP, BVAR against BVAR-SSPSV, and BVAR-SSP against BVAR-SSPSV.

The average log predictive density scores reported in Table 5 show that a rolling estimation scheme almost always yields better (lower) log scores than does a recursive estimation scheme, although sometimes by small amounts. For example, in the case of the multivariate scores for the BVAR-SSP model given in the top panel, the rolling scheme yields log scores of 8.536 at the 1 quarter horizon and 10.397 at the 1 year horizon, while the recursive scheme yields corresponding log scores of 8.780 and 10.759. In addition, using gap forms for most variables and a steady state prior usually improves log scores: given the estimation scheme (recursive or rolling), log scores are typically lower for the BVAR-SSP model than the BVAR. Continuing with the same example, the rolling BVAR has log scores of 8.600 at the 1 quarter horizon and 10.767 at the 1 year horizon (compared to the BVAR-SSP's scores of 8.536 and 10.397, respectively).

Allowing stochastic volatility offers improvements in log scores that seem especially

considerable at short horizons. In the multivariate case (top panel), the log scores of the rolling BVAR-SSPSV are 7.350 (1Q), 9.724 (2Q), 10.029 (1Y), and 12.392 (2Y), compared to the rolling BVAR-SSP model's log scores of 8.536, 10.533, 10.397, and 13.060, respectively. The improvement in the score for the full set of variables reflects improvements for most variables, although the gain is larger for short-horizon forecasts of the funds rate than other variables. For example, with the 1-quarter horizon and rolling forecasts, the BVAR-SSP and BVAR-SSPSV log scores are, respectively, 4.433 and 4.205 for GDP growth, -0.386 and -0.567 for unemployment, 2.923 and 2.811 for inflation, and 2.024 and 1.139 for the funds rate.

The p-values of the differences in average log scores reported in Table 6 indicate that the improvements in density forecasts associated with stochastic volatility are statistically meaningful at short horizons, although mixed at longer horizons. In comparing the (rolling in all cases) BVAR-SSP against the BVAR-SSPSV, the differences in average log scores are: 1.185, with a p-value of 0.000 (1Q); 0.809, with a p-value of 0.022 (2Q); 0.368, with a p-value of 0.380 (1Y); and 0.667, with a p-value of 0.100 (2Y). For the most part, the same pattern applies to each variable, except in the case of inflation forecasts, for which the differences in log scores are larger at longer horizons than shorter horizons. On the basis of this evidence, it seems reasonable to conclude that modeling stochastic volatility significantly improves the accuracy of density forecasts, although more convincingly at shorter horizons than longer horizons.

5.6 Fan chart illustration

As noted at the outset, central banks such as the Bank of Canada, Bank of England, Norges Bank, South African Reserve Bank, and Sveriges Riksbank routinely publish fan charts that provide entire forecast distributions for inflation and, in some nations, a measure of output or the policy interest rate. Recent studies such as Cogley, Morozov, and Sargent (2005) and Beechey and Osterholm (2008) have provided fan charts (for the U.K. and Australia, respectively) generated from Bayesian VARs.

To further illustrate the practical consequences of time variation in conditional volatilities, Figure 11 provides fan charts of forecasts made in the middle of 2008:Q3, using the 2008:Q3 vintage of data from the RTDSM, with a data sample ending in 2008:Q2. Following Cogley, Morozov, and Sargent (2005), the figure reports percentiles of the marginal density

for each variable at each horizon.[19] In keeping with common practice, the GDP growth and inflation forecasts are reported as forecasts of four-quarter averages. Consequently, for these variables, the probability bands widen sharply as the horizon increases from the current quarter (for which three quarters of growth and inflation entering the four-quarter average are known), to the next quarter (for which two quarters of growth and inflation entering the four-quarter average are known), and so on. In the interest of brevity, results are reported for only two of the models or methods: the BVAR-SSP and BVAR-SSPSV specifications estimated with rolling samples (most recent 20 years) of data.

For most variables considered, allowing for stochastic volatility significantly affects the fan chart estimates. As noted at the outset of this section, volatility has risen sharply in recent quarters (including 2008:Q4 and 2009:Q1, not included in the estimates underlying these fan charts). The simple rolling sample estimates that treat error variances as constant in the sample can only very gradually capture such changes. The model with stochastic volatility can more rapidly pick up the changes in error variances. As a consequence, for unemployment, inflation, and the federal funds rate, the estimated fan chart bands are wider — conveying more uncertainty surrounding the outlook — for the model with stochastic volatility (BVAR-SSPSV, rolling) than the model without (BVAR-SSP, rolling). In the case of GDP growth forecasts, though, the fan charts are quite similar for the two estimates.

6 Conclusions

Central banks and other forecasters have become increasingly interested in various aspects of density forecasts. However, recent sharp changes in macroeconomic volatility — such as the Great Moderation and the more recent sharp rise in volatility associated with greater variation in energy prices and the deep global recession — pose significant challenges to density forecasting. Accordingly, this paper examines, with real-time data, density forecasts of GDP growth, unemployment, inflation, and the federal funds rate from VAR models with stochastic volatility. The model of interest extends the steady state prior BVAR of Villani (2009) to include stochastic volatility, because, as found in some prior work and in this paper, incorporating informative priors on the steady states of the model variables often improves the accuracy of point forecasts. Motivated by results in Clark and McCracken (2008, 2010), the model also includes some variables in gap, or deviation from trend, form:

[19] These fan charts based on percentiles represent equal-tail credible sets.

unemployment less an exponential trend, inflation less the Blue Chip long-run inflation expectation, and the federal funds rate less the Blue Chip long-run inflation expectation.

The evidence presented in the paper shows that adding stochastic volatility to the BVAR with most variables in gap form and a steady state prior materially improves the real-time accuracy of point and density forecasts. The density evidence includes interval forecasts (coverage rates), histograms of probability integral transforms, time series plots and various tests applied to normal transforms of the probability integral transforms, and log predictive density scores. In the absence of stochastic volatility, models estimated with rolling samples of data are more accurate in density forecasting than models estimated recursively. But modeling stochastic volatility yields larger gains in forecast accuracy.

7 Data appendix

This appendix provides additional detail on the construction of the long-run inflation expectations series used in the VAR estimation and forecasting.

For each quarterly vintage, I construct a real-time long-run expectations series by combining exponential smoothing-based estimates for 1960:Q1-1979:Q3 with Blue Chip expectations data for 1979:Q4 through the forecast origin.

The Blue Chip data are the 6-10 year ahead (in some surveys, 5-9 year ahead) expectations of inflation in the GDP (or GNP) deflator or price index. For most of the sample, the results were published in the March and October surveys (on the 10th of each month). But for a few earlier months, the results were published in May or November. To ensure the data are available in real time, the March figures are treated as being available in the Q2 vintage, but not the Q1 vintage. Similarly, the October figures are treated as being available in the Q4 vintage, but not the Q3 vintage.

The exponentially smoothed series are constructed in real time, for each quarterly vintage t, as follows. (1) Initialize the filter with the mean rate of price increase for 1953-59, using that vintage's price series. If data are not available for that 1953-59 period, use the same mean from the most recent vintage with data available for that period. The mean becomes the exponentially smoothed estimate for period 1959:4. Use exponential smoothing with a smoothing coefficient of .05, to estimate trend inflation from 1960:1 through t-1. Define the value of long-run expected inflation for period t as the exponentially smoothed trend estimated with data through t-1. Because the first vintage of price data is 1965:Q4, this approach begins with that vintage. To make things as real time as possible, estimates of trend inflation for 1960:Q31-1965:Q3 are also those obtained by exponential smoothing of the 1965:Q4 vintage of data. Note, however, that applying exponential smoothing to the historical data of the 2008:Q4 vintage yields a very similar trend estimate.

For each vintage t starting in 1985:Q1, a time series on inflation expectations is compiled as follows. For 1960:Q1 through 1979:Q3, the expectation is estimated with the real time exponentially smoothed series. For 1979:Q4 through t, the expectation is based on Blue Chip. For odd-numbered quarters before period t (in vintage t), I linearly interpolate between quarters 2 and 4. When period t of vintage t is odd-numbered, the expectation is simply the Blue Chip value from period t-1.

References

Adolfson, Malin, Michael K. Andersson, Jesper Linde, Mattias Villani, and Anders Vredin (2007), "Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks," *International Journal of Central Banking*, December, 111-144.

Adolfson, Malin, Jesper Linde, and Mattias Villani (2005), "Forecasting Performance of an Open Economy Dynamic Stochastic General Equilibrium Model," Sveriges Riksbank Working Paper No. 190.

Amisano, Gianni, and Raffaella Giacomini (2007), "Comparing Density Forecasts via Weighted Likelihood Ratio Tests," *Journal of Business and Economic Statistics* 25, 177-190.

Beechey, Meredith, and Par Osterholm (2008), "A Bayesian Vector Autoregressive Model with Informative Steady-State Priors for the Australian Economy," *Economic Record* 84, 449-465.

Berkowitz, Jeremy (2001), "Testing Density Forecasts, With Applications to Risk Management," *Journal of Business and Economic Statistics* 19, 465-474.

Christoffersen, Peter F. (1998), "Evaluating Interval Forecasts," *International Economic Review* 39, 841-862.

Christoffersen, Peter, and Stefano Mazzotta (2005), "The Accuracy of Density Forecasts from Foreign Exchange Options," *Journal of Financial Econometrics* 3, 578-605.

Clark, Todd E. and Michael W. McCracken (2010), "Averaging Forecasts from VARs with Uncertain Instabilities," *Journal of Applied Econometrics*, forthcoming.

Clark, Todd E. and Michael W. McCracken (2008), "Forecasting with Small Macroeconomic VARs in the Presence of Instability," in *Forecasting in the Presence of Structural Breaks and Model Uncertainty*, Rapach DE, Wohar ME (eds). Emerald Publishing.

Clements, Michael P. (2004), "Evaluating the Bank of England Density Forecasts of Inflation," *Economic Journal* 114, 844-866.

Cogley Timothy, and Thomas J. Sargent (2005), "Drifts and Volatilities: Monetary Policies and Outcomes in the Post-World War II U.S.," *Review of Economic Dynamics* 8, 262-302.

Cogley, Timothy, Sergei Morozov, and Thomas J. Sargent (2005), "Bayesian Fan Charts for U.K. Inflation: Forecasting and Sources of Uncertainty in an Evolving Monetary System," *Journal of Economic Dynamics and Control* 29, 1893-1925.

Croushore, Dean and Tom Stark (2001), "A Real-Time Data Set for Macroeconomists," *Journal of Econometrics* 105, 111-30.

Croushore, Dean (2006), "Forecasting with Real-Time Macroeconomic Data," *Handbook of Economic Forecasting*, G. Elliott, C. Granger, and A. Timmermann (Eds.), North Holland: Amsterdam.

D'Agostino, Antonella, Luca Gambetti, and Domenico Giannone (2008), "Macroeconomic Forecasting and Structural Change," manuscript, European Central Bank.

Diebold, Francis X., Todd A. Gunther, and Anthony S. Tay (1998), "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review* 39, 863-883.

Faust, Jon, and Jonathan H. Wright (2007), "Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset," NBER Working Paper #13397.

Geweke, John, and Gianni Amisano (2008), "Comparing and Evaluating Bayesian Predictive Distributions of Asset Returns," manuscript, University of Iowa.

Giordani, Paolo, and Mattias Villani (2008), "Forecasting Macroeconomic Time Series with Locally Adaptive Signal Extraction," manuscript, Sveriges Riksbank.

Hong, Yongmiao, Haitao Li, and Feng Zhao (2004), "Out-of-Sample Performance of Discrete-Time Spot Interest Rate Models," *Journal of Business and Economic Statistics* 22, 457-473.

Jacquier, Eric, Nicholas G. Polson, and Peter E. Rossi, (1994), "Bayesian Analysis of Stochastic Volatility Models," *Journal of Business and Economic Statistics* 12, 371418.

Kadiyala, K. Rao and Sune Karlsson (1997), "Numerical Methods for Estimation and Inference in Bayesian VAR Models," *Journal of Applied Econometrics* 12, 99-132.

Kozicki, Sharon and Peter A. Tinsley (2001a), "Shifting Endpoints in the Term Structure of Interest Rates," *Journal of Monetary Economics* 47, 613-652.

Kozicki, Sharon and Peter A. Tinsley (2001b), "Term Structure Views of Monetary Policy under Alternative Models of Agent Expectations," *Journal of Economic Dynamics and Control* 25, 149-84.

Jore, Anne Sofie, James Mitchell, and Shaun P. Vahey (2009), "Combining Forecast Densities from VARS with Uncertain Instabilities," *Journal of Applied Econometrics*, forthcoming.

Osterholm, Par (2008), "Can Forecasting Performance Be Improved by Considering the Steady State? An Application to Swedish Inflation and Interest Rate," *Journal of Forecasting* 27, 41-51.

Primiceri, Giorgio (2005), "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies* 72, 821-852

Robertson, John C., Ellis W. Tallman, and Charles H. Whiteman (2005), "Forecasting Using Relative Entropy," *Journal of Money, Credit and Banking* 37, 383-401.

Romer, Christina D., and David H. Romer (2000), "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review* 90, 429-457.

Sims, Christopher A. (2002), "The Role of Models and Probabilities in the Monetary Policy Process," *Brookings Papers on Economic Activity* 2, 1-40.

Sims, Christopher A., and Tao Zha (2006), "Were There Regime Switches in US Monetary Policy?" *American Economic Review* 96, 54-81.

Stock, James H., and Mark W. Watson (2007), "Has U.S. Inflation Become Harder to

Forecast?" *Journal of Money, Credit, and Banking* 39, 3-33.

Villani, Mattias (2009), "Steady-State Priors for Vector Autoregressions," *Journal of Applied Econometrics* 24, 630-650.

Table 1. Real-Time Mean Forecast Errors, 1985-2008Q3

	$h = 1Q$	$h = 2Q$	$h = 1Y$	$h = 2Y$
	univariate models			
GDP growth	-0.307	-0.352	-0.331	-0.409
unemployment	0.002	-0.000	-0.026	-0.099
GDP inflation	-0.001	-0.030	-0.023	-0.044
fed funds rate	-0.045	-0.117	-0.218	-0.323
	BVAR, recursive			
GDP growth	-1.040	-1.076	-0.930	-0.562
unemployment	0.039	0.084	0.160	0.159
GDP inflation	-0.227	-0.497	-0.556	-1.203
fed funds rate	-0.178	-0.369	-0.693	-1.221
	BVAR, rolling			
GDP growth	-0.355	-0.323	-0.213	0.126
unemployment	0.032	0.061	0.091	0.009
GDP inflation	-0.197	-0.432	-0.499	-1.093
fed funds rate	-0.160	-0.338	-0.668	-1.249
	BVAR-SSP, recursive			
GDP growth	-0.703	-0.675	-0.540	-0.230
unemployment	0.019	0.037	0.049	-0.051
GDP inflation	-0.098	-0.294	-0.335	-0.770
fed funds rate	-0.167	-0.347	-0.645	-1.091
	BVAR-SSP, rolling			
GDP growth	-0.192	-0.125	-0.064	0.054
unemployment	-0.007	-0.019	-0.064	-0.205
GDP inflation	-0.011	-0.129	-0.161	-0.446
fed funds rate	-0.087	-0.189	-0.384	-0.709
	BVAR-SSPSV, recursive			
GDP growth	-0.389	-0.388	-0.300	-0.127
unemployment	0.006	0.012	0.010	-0.071
GDP inflation	-0.069	-0.217	-0.241	-0.557
fed funds rate	-0.110	-0.250	-0.503	-0.871
	BVAR-SSPSV, rolling			
GDP growth	-0.249	-0.211	-0.147	-0.015
unemployment	0.007	0.009	0.001	-0.071
GDP inflation	-0.003	-0.095	-0.113	-0.320
fed funds rate	-0.084	-0.198	-0.411	-0.726

Notes:
1. The entries in the first panels are mean forecast errors (actuals less forecasts), for variables defined in annualized percentage points.
2. The forecast errors are calculated using the second–available (real–time) estimates of growth and inflation as the actual data, and currently available measures of unemployment and the federal funds rate as actuals.
3. In each quarter t from 1985:Q1 through 2008:Q3, vintage t data (which generally end in $t-1$) are used to form forecasts for periods t ($h = 1Q$), $t+1$ ($h = 2Q$), $t+3$ ($h = 1Y$), and $t+7$ ($h = 2Y$). The forecasts of GDP growth and inflation for the $h = 1Y$ and $h = 2Y$ horizons correspond to annual percent changes: average growth and average inflation from t through $t+3$ and $t+4$ through $t+7$, respectively.

Table 2. Real-Time Forecast RMSEs, 1985-2008Q3
(RMSEs for univariate models, RMSE ratios in all others)

	$h = 1Q$	$h = 2Q$	$h = 1Y$	$h = 2Y$
	univariate models			
GDP growth	1.786	1.828	1.319	1.444
unemployment	0.173	0.313	0.574	1.055
GDP inflation	0.996	1.046	0.683	0.924
fed funds rate	0.396	0.813	1.489	2.390
	BVAR, recursive			
GDP growth	1.240	1.220	1.216	0.925
unemployment	1.047	1.030	0.989	0.722
GDP inflation	1.068	1.091	1.281	1.586
fed funds rate	1.216	1.122	1.047	0.994
	BVAR, rolling			
GDP growth	1.148	1.135	1.087	0.909
unemployment	1.053	1.014	0.932	0.662
GDP inflation	1.081	1.154	1.432	1.805
fed funds rate	1.204	1.083	1.021	1.001
	BVAR-SSP, recursive			
GDP growth	1.157	1.129	1.059	0.896
unemployment	1.048	1.014	0.949	0.728
GDP inflation	1.019	0.978	1.034	1.107
fed funds rate	1.203	1.106	0.997	0.888
	BVAR-SSP, rolling			
GDP growth	1.091	1.049	0.954	0.844
unemployment	1.013	0.960	0.877	0.686
GDP inflation	1.013	0.980	1.013	1.078
fed funds rate	1.134	1.006	0.902	0.813
	BVAR-SSPSV, recursive			
GDP growth	1.074	1.060	0.984	0.885
unemployment	0.982	0.940	0.899	0.735
GDP inflation	1.002	0.974	0.975	1.002
fed funds rate	1.008	0.947	0.914	0.855
	BVAR-SSPSV, rolling			
GDP growth	1.057	1.031	0.939	0.860
unemployment	0.985	0.931	0.859	0.696
GDP inflation	1.004	0.987	0.989	0.996
fed funds rate	0.982	0.907	0.862	0.804

Notes:
1. See the notes to Table 1.
2. The entries in the first panel are RMSEs, for variables defined in annualized percentage points. All other entries are RMSE ratios, for the indicated model relative to the corresponding univariate model.

Table 3. Real-Time Forecast Coverage Rates, 1985-2008Q3
(Frequencies of Actual Outcomes Falling Inside 70% Intervals)

	$h = 1Q$	$h = 2Q$	$h = 1Y$	$h = 2Y$
	BVAR, recursive			
GDP growth	0.863 (0.000)	0.883 (0.000)	0.674 (0.594)	0.852 (0.000)
unemployment	0.874 (0.000)	0.851 (0.000)	0.739 (0.393)	0.682 (0.714)
GDP inflation	0.811 (0.006)	0.830 (0.001)	0.870 (0.000)	0.795 (0.026)
fed funds rate	0.947 (0.000)	0.915 (0.000)	0.837 (0.000)	0.784 (0.055)
	BVAR, rolling			
GDP growth	0.821 (0.002)	0.851 (0.000)	0.707 (0.891)	0.852 (0.000)
unemployment	0.821 (0.002)	0.819 (0.003)	0.717 (0.711)	0.636 (0.215)
GDP inflation	0.758 (0.190)	0.755 (0.212)	0.783 (0.055)	0.659 (0.418)
fed funds rate	0.926 (0.000)	0.883 (0.000)	0.793 (0.027)	0.705 (0.926)
	BVAR-SSP, recursive			
GDP growth	0.884 (0.000)	0.894 (0.000)	0.804 (0.012)	0.920 (0.000)
unemployment	0.863 (0.000)	0.840 (0.000)	0.783 (0.055)	0.784 (0.055)
GDP inflation	0.842 (0.000)	0.894 (0.000)	0.913 (0.000)	0.943 (0.000)
fed funds rate	0.947 (0.000)	0.904 (0.000)	0.826 (0.001)	0.739 (0.409)
	BVAR-SSP, rolling			
GDP growth	0.853 (0.000)	0.894 (0.000)	0.793 (0.027)	0.909 (0.000)
unemployment	0.842 (0.000)	0.819 (0.003)	0.783 (0.055)	0.818 (0.004)
GDP inflation	0.737 (0.417)	0.819 (0.003)	0.804 (0.012)	0.830 (0.001)
fed funds rate	0.905 (0.000)	0.904 (0.000)	0.837 (0.000)	0.727 (0.566)
	BVAR-SSPSV, recursive			
GDP growth	0.705 (0.911)	0.745 (0.320)	0.696 (0.928)	0.739 (0.409)
unemployment	0.705 (0.911)	0.702 (0.964)	0.674 (0.594)	0.591 (0.037)
GDP inflation	0.695 (0.912)	0.745 (0.320)	0.848 (0.000)	0.773 (0.104)
fed funds rate	0.800 (0.015)	0.670 (0.539)	0.576 (0.016)	0.568 (0.013)
	BVAR-SSPSV, rolling			
GDP growth	0.695 (0.912)	0.755 (0.212)	0.717 (0.711)	0.727 (0.566)
unemployment	0.737 (0.417)	0.734 (0.455)	0.696 (0.928)	0.693 (0.890)
GDP inflation	0.705 (0.911)	0.723 (0.612)	0.772 (0.101)	0.784 (0.055)
fed funds rate	0.789 (0.033)	0.734 (0.455)	0.728 (0.542)	0.625 (0.146)

Notes:
1. See the notes to Table 1.
2. The table reports the frequencies with which actual outcomes fall within 70 percent bands computed from the posterior distribution of forecasts.
3. The table includes in parentheses p-values for the null of correct coverage (empirical = nominal rate of 70 percent), based on t-statistics (based on OLS variances at the 1-step horizon and autocorrelation-consistent variances at other horizons).

Table 4. Tests of Normalized Errors of 1-Step Ahead Real-Time Forecasts, 1985-2008Q3

	variance (p-value)	mean (p-value)	AR(1) coef. (p-value)	LR test p-value
	BVAR, recursive			
GDP growth	0.506 (.000)	-0.334 (0.001)	0.337 (0.001)	0.000
unemployment	0.527 (.000)	0.162 (0.151)	0.401 (0.000)	0.000
GDP inflation	0.672 (.001)	-0.169 (0.004)	-0.204 (0.053)	0.002
fed funds rate	0.200 (.000)	-0.165 (0.022)	0.667 (0.000)	0.000
	BVAR, rolling			
GDP growth	0.534 (.000)	-0.107 (0.320)	0.266 (0.010)	0.000
unemployment	0.675 (.070)	0.140 (0.254)	0.381 (0.000)	0.000
GDP inflation	0.901 (.512)	-0.103 (0.228)	-0.143 (0.181)	0.332
fed funds rate	0.400 (.000)	-0.192 (0.026)	0.537 (0.000)	0.000
	BVAR-SSP, recursive			
GDP growth	0.442 (0.000)	-0.226 (0.021)	0.310 (0.002)	0.000
unemployment	0.556 (0.001)	0.092 (0.426)	0.395 (0.000)	0.000
GDP inflation	0.637 (0.000)	-0.071 (0.230)	-0.199 (0.060)	0.005
fed funds rate	0.206 (0.000)	-0.162 (0.030)	0.673 (0.000)	0.000
	BVAR-SSP, rolling			
GDP growth	0.516 (0.000)	-0.053 (0.587)	0.202 (0.054)	0.000
unemployment	0.650 (0.040)	-0.002 (0.985)	0.336 (0.002)	0.001
GDP inflation	0.885 (0.476)	0.030 (0.705)	-0.153 (0.153)	0.418
fed funds rate	0.406 (0.000)	-0.137 (0.130)	0.544 (0.000)	0.000
	BVAR-SSPSV, recursive			
GDP growth	0.852 (0.223)	-0.158 (0.215)	0.189 (0.070)	0.063
unemployment	0.921 (0.614)	0.024 (0.856)	0.203 (0.059)	0.269
GDP inflation	1.030 (0.805)	-0.089 (0.265)	-0.124 (0.241)	0.540
fed funds rate	0.848 (0.361)	-0.225 (0.097)	0.494 (0.000)	0.000
	BVAR-SSPSV, rolling			
GDP growth	0.851 (0.186)	-0.086 (0.482)	0.181 (0.082)	0.168
unemployment	0.879 (0.455)	0.028 (0.832)	0.267 (0.012)	0.071
GDP inflation	1.005 (0.965)	-0.003 (0.969)	-0.116 (0.274)	0.751
fed funds rate	0.760 (0.201)	-0.184 (0.125)	0.482 (0.000)	0.000

Notes:
1. See the notes to Table 1.
2. The normalized forecast error is defined as $\Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of the one-step ahead forecast error and Φ^{-1} is the inverse of the standard normal distribution function.
3. The first column reports the estimated variance of the normalized error, along with a p-value for a test of the null hypothesis of a variance equal to 1 (computed by a linear regression of the squared error on a constant, using a Newey-West variance). The second column reports the mean of the normalized error, along with a p-value for a test of the null of a mean of zero (using a Newey-West variance). The third column reports the AR(1) coefficient and its p-value, obtained by estimating a linear model with an intercept. The final column reports the p-value of Berkowitz's (2001) likelihood ratio test for the joint null of a zero mean, unity variance, and no (AR(1)) serial correlation.

Table 5. Real-Time Forecast Average Log Scores, 1985-2008Q3

	$h = 1Q$	$h = 2Q$	$h = 1Y$	$h = 2Y$
	all four variables			
BVAR, recursive	8.860	11.005	11.092	13.734
BVAR, rolling	8.600	10.731	10.767	13.582
BVAR-SSP, recursive	8.780	10.823	10.759	13.322
BVAR-SSP, rolling	8.536	10.533	10.397	13.060
BVAR-SSPSV, recursive	7.343	9.957	10.515	12.959
BVAR-SSPSV, rolling	7.350	9.724	10.029	12.392
	GDP growth			
BVAR, recursive	4.628	4.648	3.812	3.732
BVAR, rolling	4.470	4.496	3.520	3.632
BVAR-SSP, recursive	4.566	4.584	3.599	3.769
BVAR-SSP, rolling	4.433	4.443	3.386	3.594
BVAR-SSPSV, recursive	4.245	4.334	3.514	3.517
BVAR-SSPSV, rolling	4.205	4.275	3.411	3.482
	unemployment			
BVAR, recursive	-0.360	0.675	1.733	2.319
BVAR, rolling	-0.344	0.629	1.602	2.201
BVAR-SSP, recursive	-0.368	0.655	1.681	2.423
BVAR-SSP, rolling	-0.386	0.589	1.550	2.329
BVAR-SSPSV, recursive	-0.605	0.504	1.600	2.371
BVAR-SSPSV, rolling	-0.567	0.536	1.565	2.329
	inflation			
BVAR, recursive	3.014	3.272	2.834	3.739
BVAR, rolling	3.008	3.222	2.767	3.636
BVAR-SSP, recursive	2.944	3.132	2.604	3.360
BVAR-SSP, rolling	2.923	3.062	2.469	3.170
BVAR-SSPSV, recursive	2.810	2.923	2.257	3.001
BVAR-SSPSV, rolling	2.811	2.928	2.261	2.938
	federal funds rate			
BVAR, recursive	2.224	3.107	3.935	4.662
BVAR, rolling	2.058	3.052	3.970	4.753
BVAR-SSP, recursive	2.189	3.051	3.818	4.428
BVAR-SSP, rolling	2.024	2.988	3.811	4.409
BVAR-SSPSV, recursive	1.090	2.528	3.952	4.634
BVAR-SSPSV, rolling	1.139	2.440	3.649	4.433

Notes:
1. See the notes to Table 1.
2. The table entries are average values of log predictive density scores, computed with the Gaussian (quadratic) formula given in Adolfson, Linde, and Villani (2005), under which a lower score implies a better model.

Table 6. Amisano-Giacomini Test Applied to Average Log Scores of Real-Time Forecasts, 1985-2008Q3

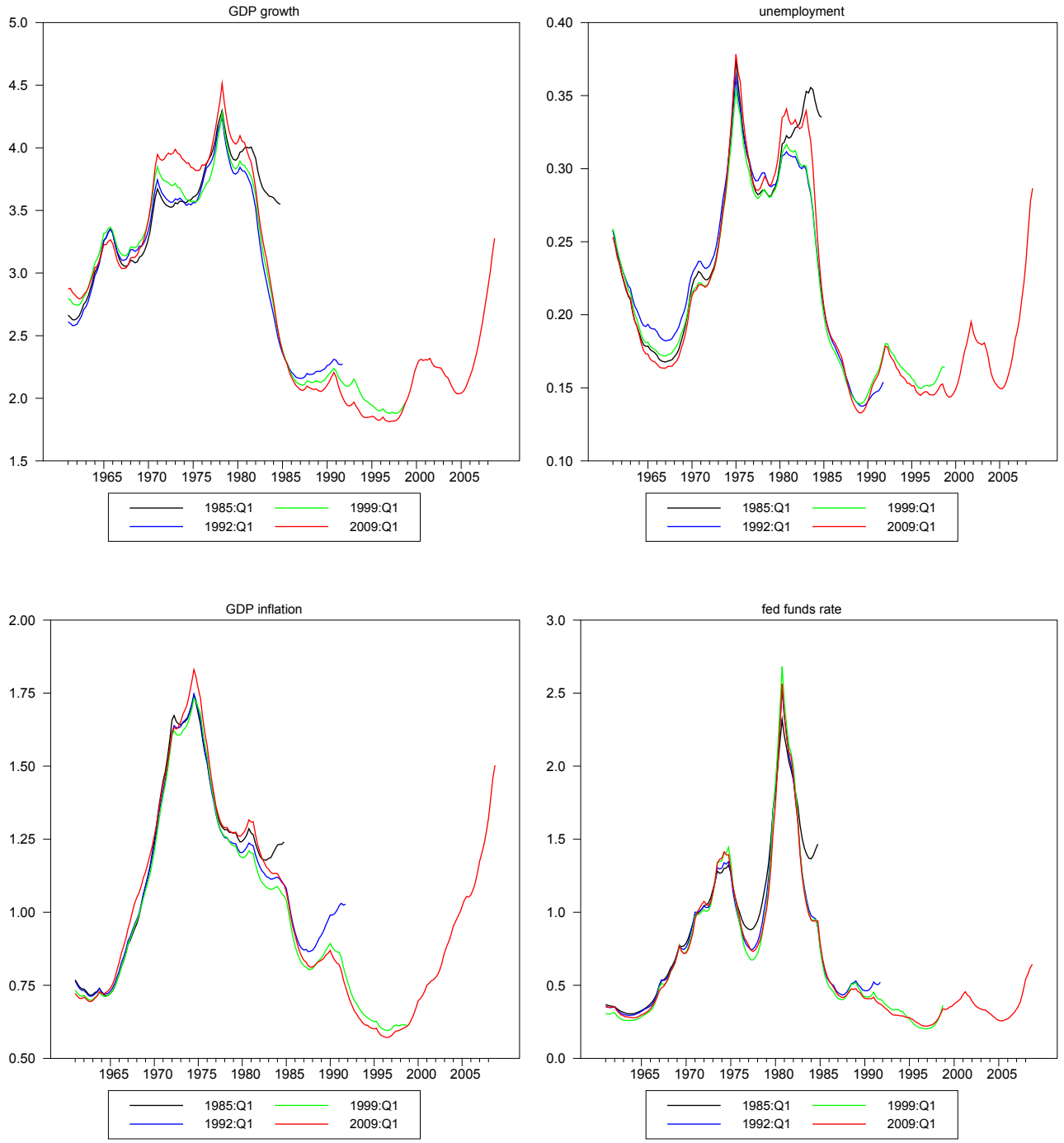
(Mean Differences in Log Scores, with p-values)

	$h = 1Q$	$h = 2Q$	$h = 1Y$	$h = 2Y$
	all four variables			
BVAR vs. BVAR-SSP	0.065 (0.221)	0.198 (0.031)	0.370 (0.053)	0.522 (0.066)
BVAR vs. BVAR-SSPSV	1.250 (0.000)	1.007 (0.007)	0.738 (0.110)	1.189 (0.032)
BVAR-SSP vs. BVAR-SSPSV	1.185 (0.000)	0.809 (0.022)	0.368 (0.380)	0.667 (0.100)
	GDP growth			
BVAR vs. BVAR-SSP	0.038 (0.084)	0.054 (0.018)	0.134 (0.043)	0.038 (0.458)
BVAR vs. BVAR-SSPSV	0.266 (0.000)	0.221 (0.010)	0.109 (0.311)	0.149 (0.079)
BVAR-SSP vs. BVAR-SSPSV	0.228 (0.000)	0.168 (0.071)	-0.025 (0.846)	0.112 (0.318)
	unemployment			
BVAR vs. BVAR-SSP	0.042 (0.193)	0.040 (0.586)	0.051 (0.770)	-0.129 (0.412)
BVAR vs. BVAR-SSPSV	0.223 (0.000)	0.093 (0.321)	0.037 (0.803)	-0.128 (0.297)
BVAR-SSP vs. BVAR-SSPSV	0.181 (0.002)	0.053 (0.607)	-0.014 (0.923)	0.000 (0.997)
	inflation			
BVAR vs. BVAR-SSP	0.085 (0.006)	0.161 (0.000)	0.297 (0.000)	0.466 (0.003)
BVAR vs. BVAR-SSPSV	0.197 (0.001)	0.295 (0.000)	0.506 (0.000)	0.699 (0.007)
BVAR-SSP vs. BVAR-SSPSV	0.112 (0.025)	0.134 (0.007)	0.209 (0.013)	0.233 (0.046)
	federal funds rate			
BVAR vs. BVAR-SSP	0.035 (0.024)	0.064 (0.006)	0.159 (0.000)	0.344 (0.000)
BVAR vs. BVAR-SSPSV	0.919 (0.000)	0.612 (0.003)	0.321 (0.180)	0.320 (0.189)
BVAR-SSP vs. BVAR-SSPSV	0.884 (0.000)	0.548 (0.008)	0.162 (0.525)	-0.024 (0.931)

Notes:
1. See the notes to Table 1.
2. The table entries are differences in average log predictive density scores and p-values from Amisano and Giacomini (2007) tests of equal average scores. The tests and p-values are computed by regressions of differences in log scores (time series) on a constant, using the Newey-West estimator of the variance of the regression constant. All of the results in the table are based on forecasts from models estimated with rolling samples of data.

Figure 1. Posterior Means of Residual Standard Deviations, Final Vintage Data

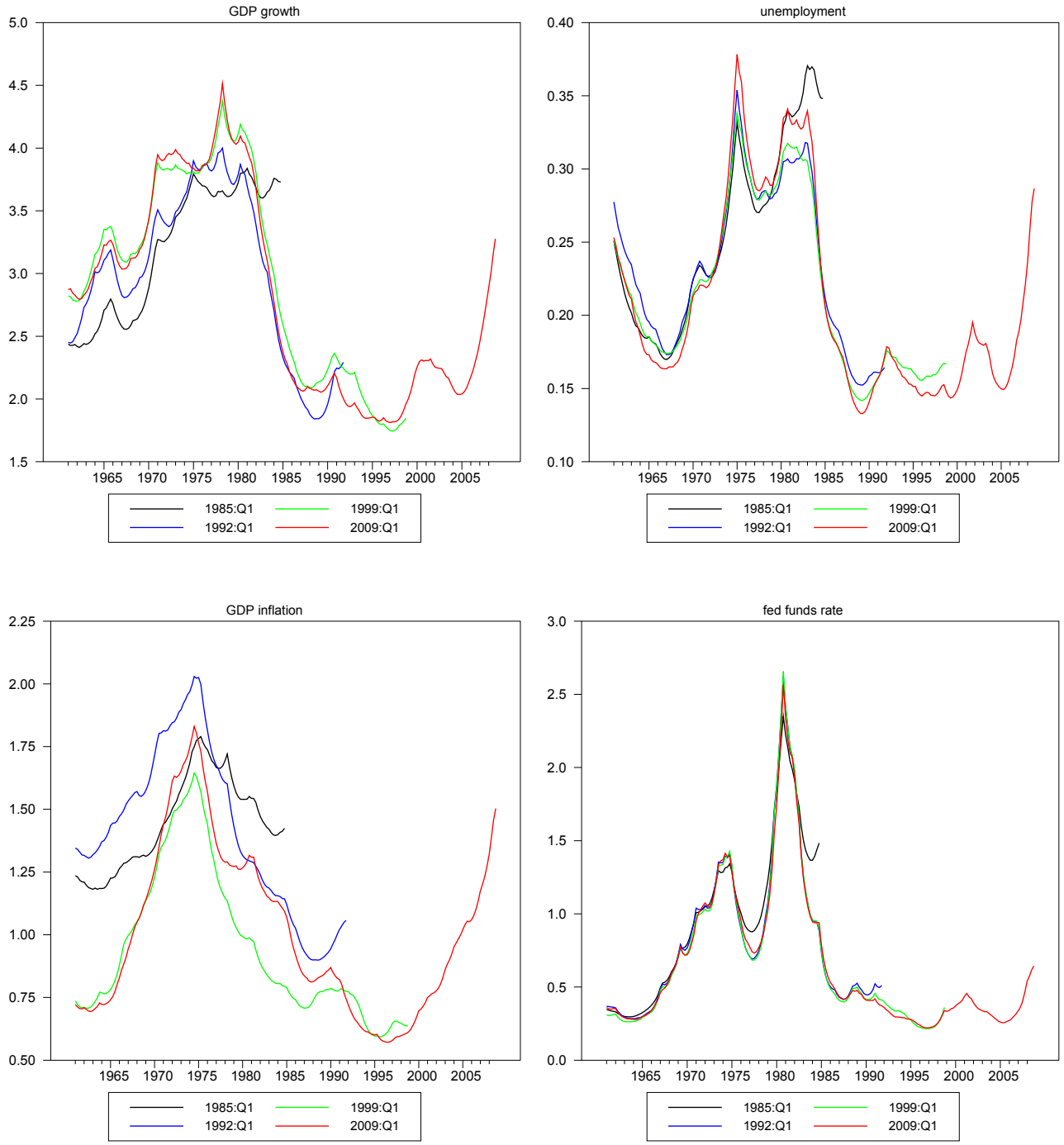
BVAR-SSPSV, recursive



Notes: The figure reports posterior means of the time series of estimates of the residual variances in the BVAR-SSPSV model (recursive), estimated at various points in time with the 2009:Q1 vintage of data, which includes data through 2008:Q4. The dates given for each line (1985:Q1, 1992:Q1, etc.) correspond to pseudo-vintage dates; at each of these points in time, the model was estimated with data through the prior quarter.

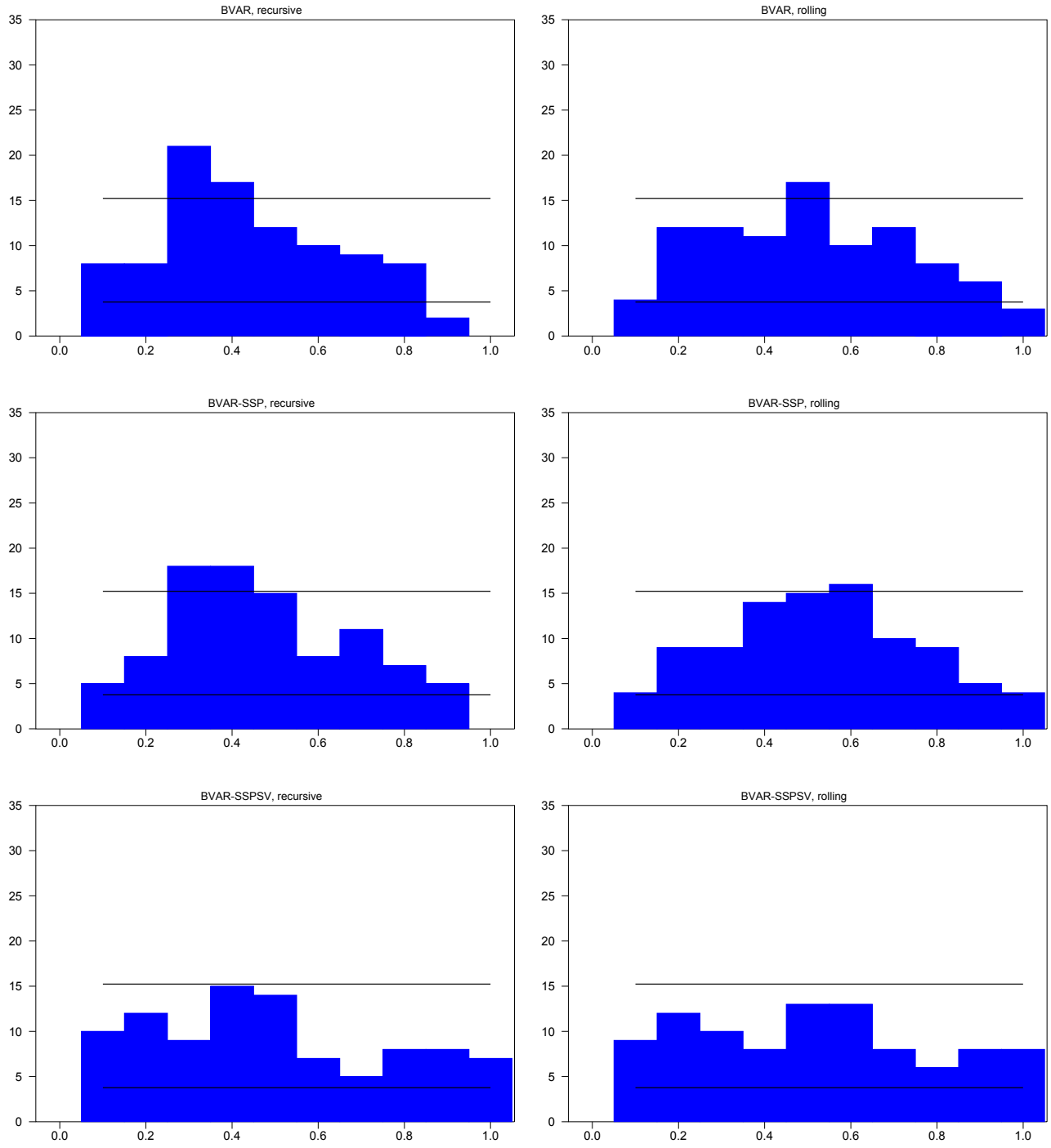
Figure 2. Posterior Means of Residual Standard Deviations, Real-Time Vintages

BVAR-SSPSV, recursive



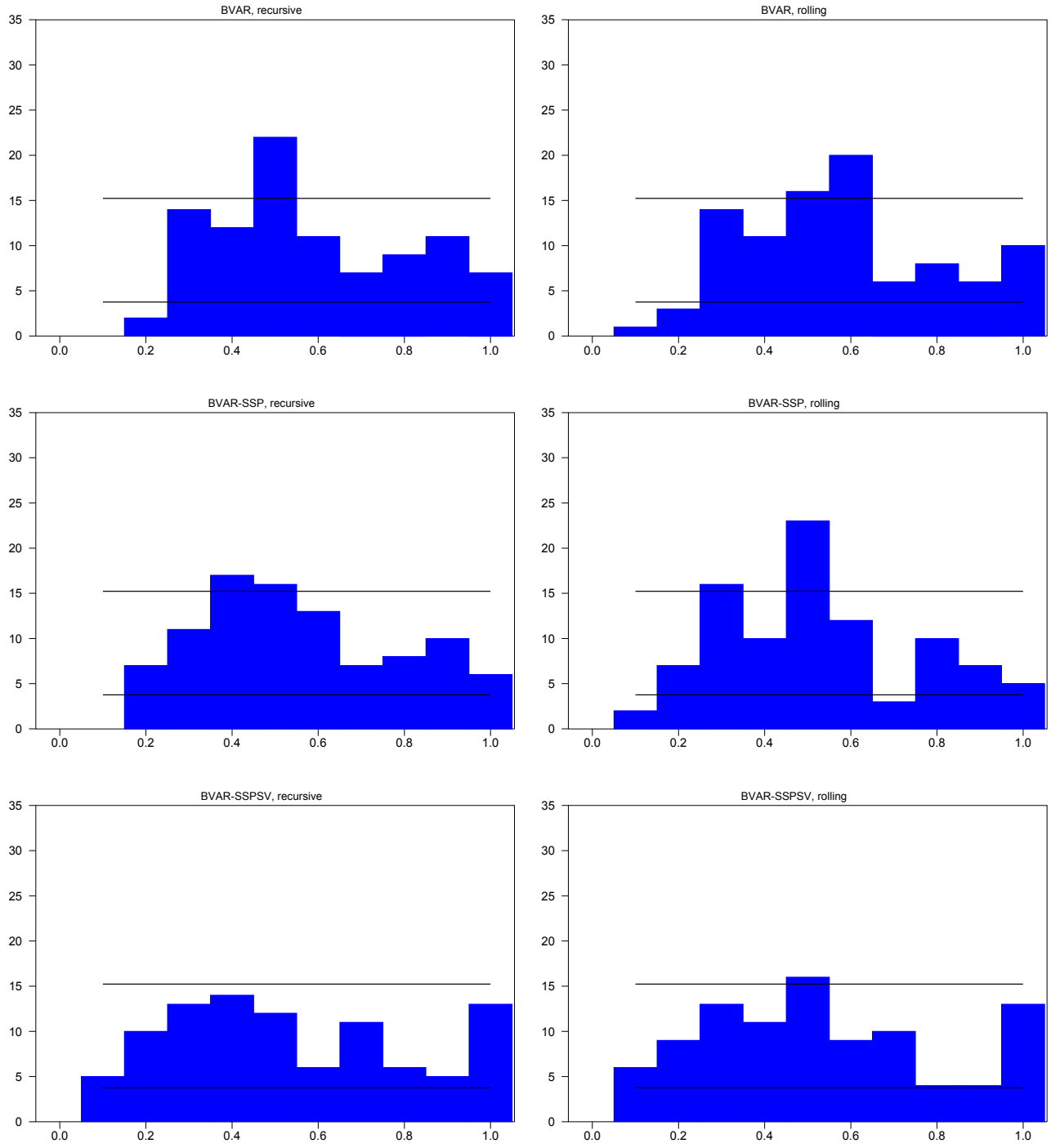
Notes: The figure reports posterior means of the time series of estimates of the residual variances in the BVAR-SSPSV model (recursive), estimated at various points in time with the vintage of data indicated. The dates given for each line (1985:Q1, 1992:Q1, etc.) correspond to the dates of the data vintages; at each of these points in time, the model was estimated with the indicated vintage, using data through the prior quarter.

Figure 3. PIT histogram for GDP growth
horizon = 1Q



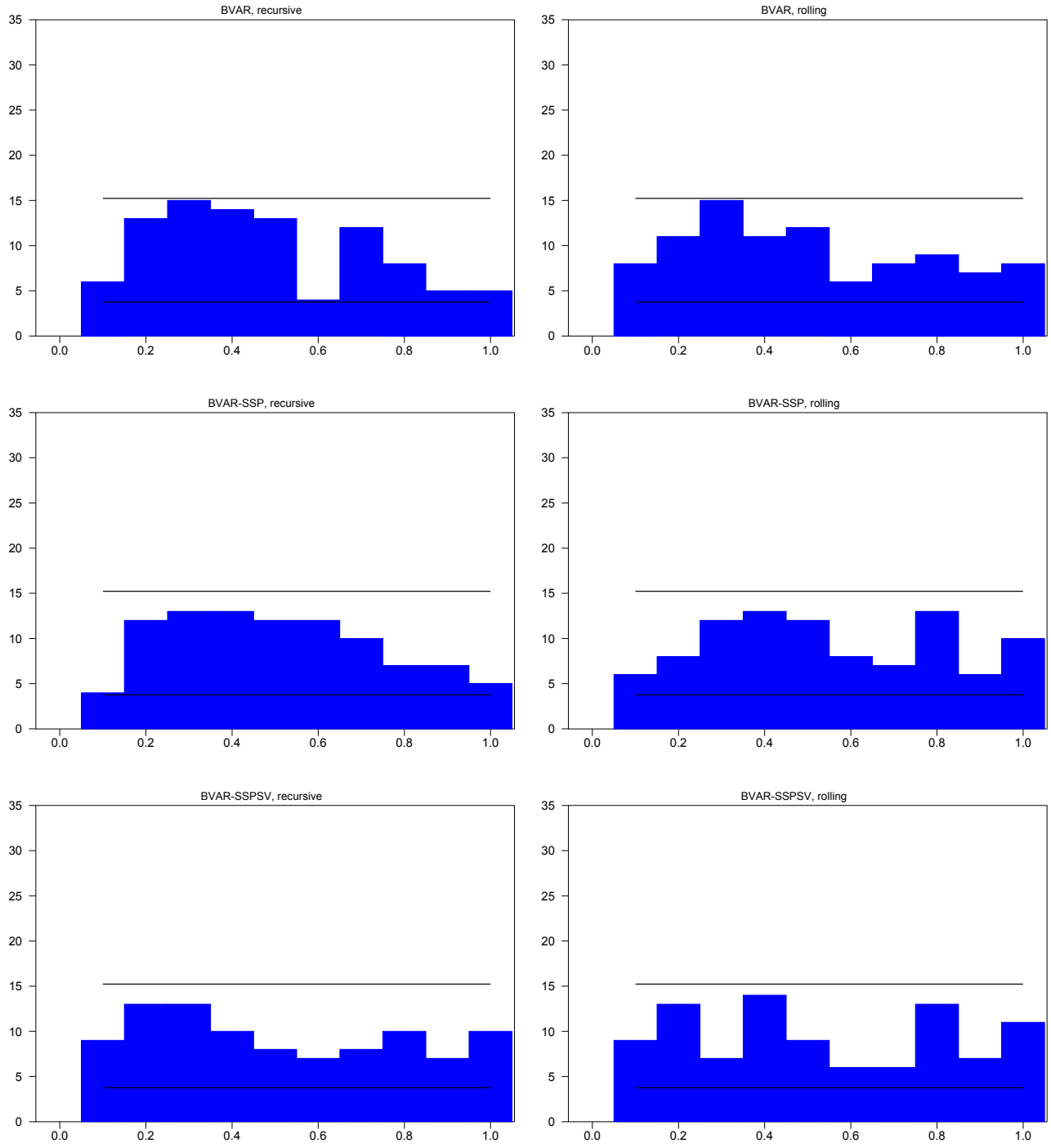
Note: The histograms shown are decile counts of the PITs transforms based on real-time forecasts, with 95 percent intervals (intended only as a rough guide) estimated under the binomial distribution.

Figure 4. PIT histogram for unemployment
horizon = 1Q



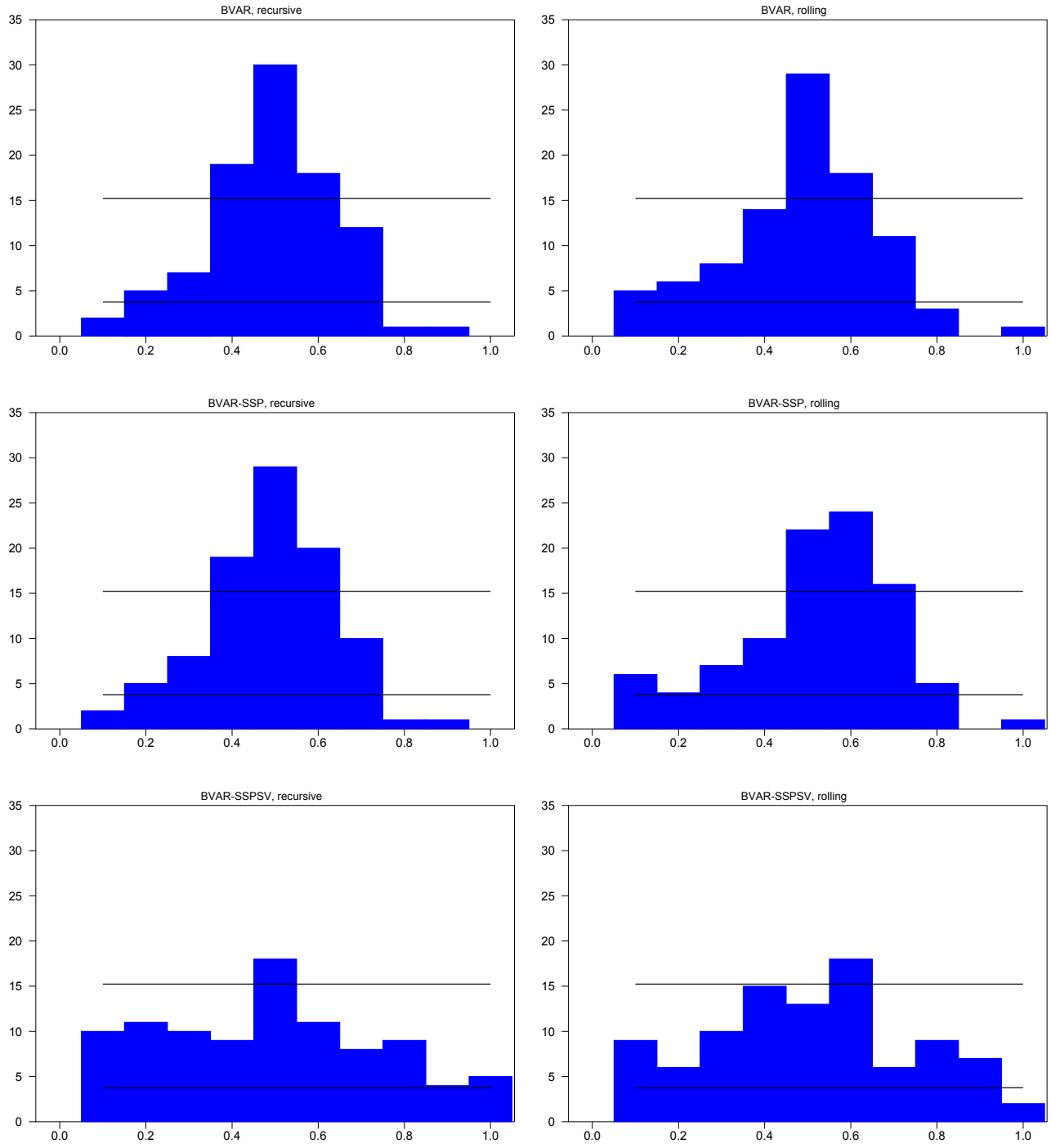
Note: The histograms shown are decile counts of the PITs transforms based on real-time forecasts, with 95 percent intervals (intended only as a rough guide) estimated under the binomial distribution.

Figure 5. PIT histogram for GDP inflation
horizon = 1Q



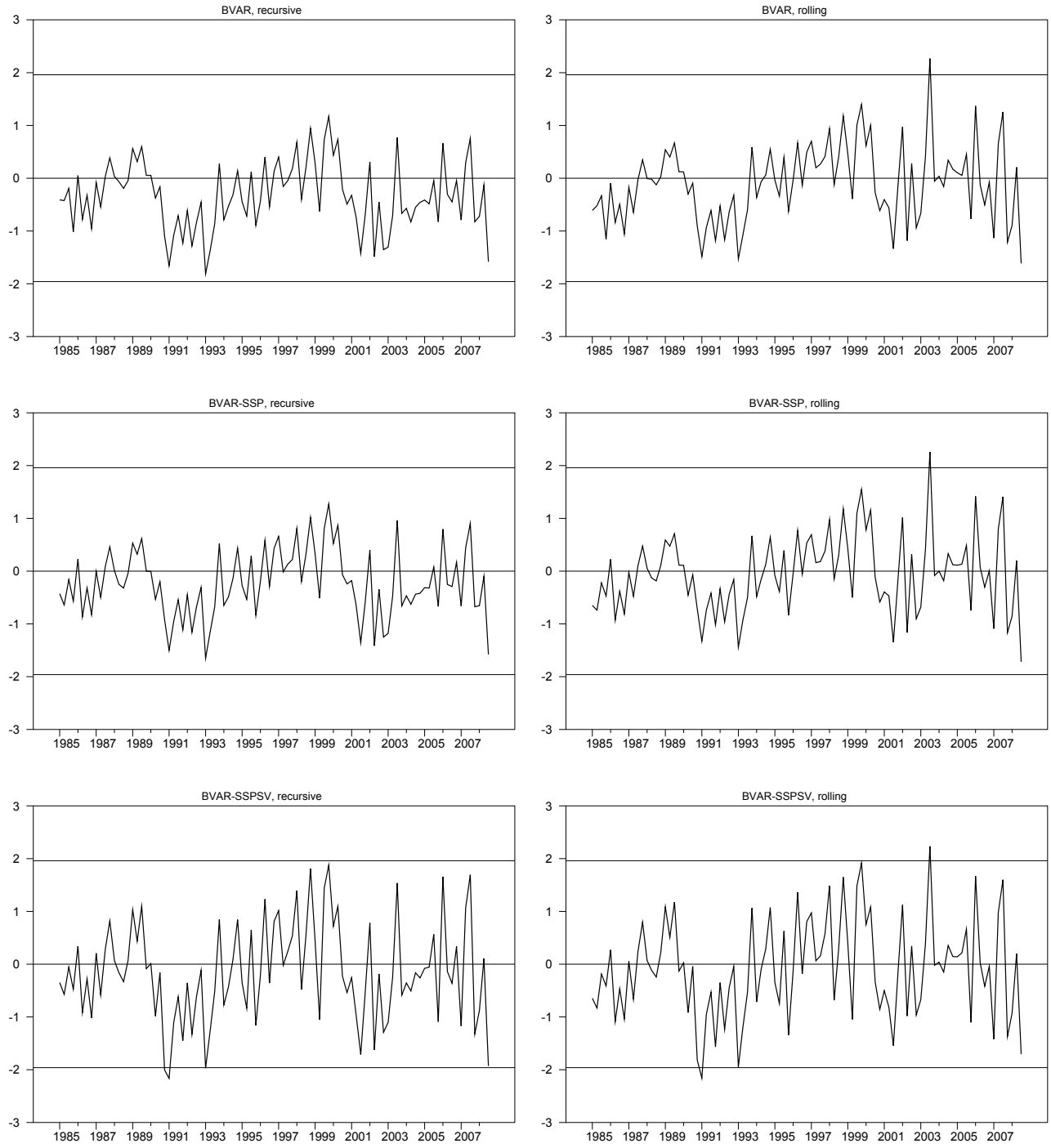
Note: The histograms shown are decile counts of the PITs transforms based on real-time forecasts, with 95 percent intervals (intended only as a rough guide) estimated under the binomial distribution.

Figure 6. PIT histogram for fed funds rate
horizon = 1Q



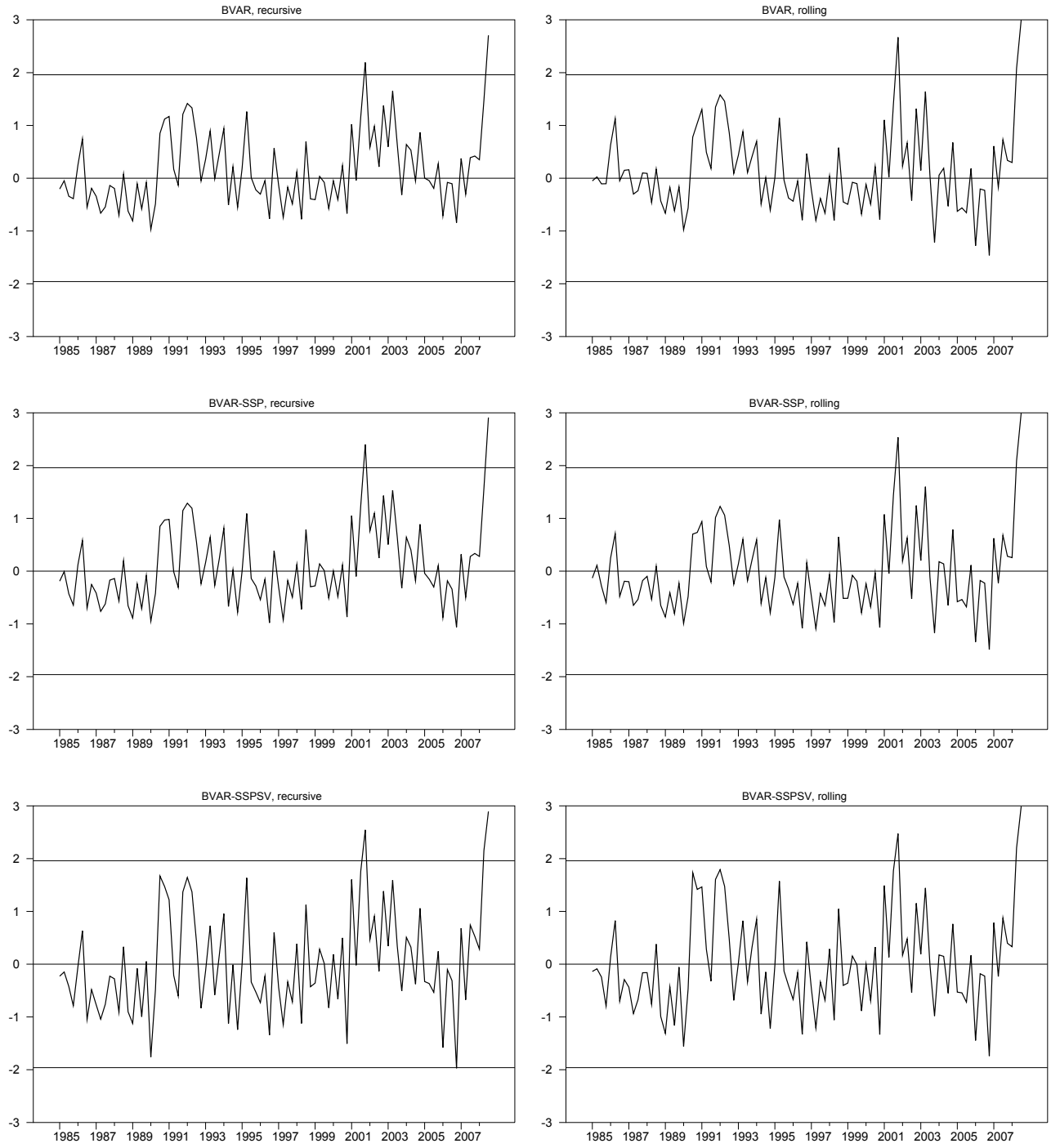
Note: The histograms shown are decile counts of the PITs transforms based on real-time forecasts, with 95 percent intervals (intended only as a rough guide) estimated under the binomial distribution.

Figure 7. Normalized Forecast Errors for GDP growth
horizon = 1Q



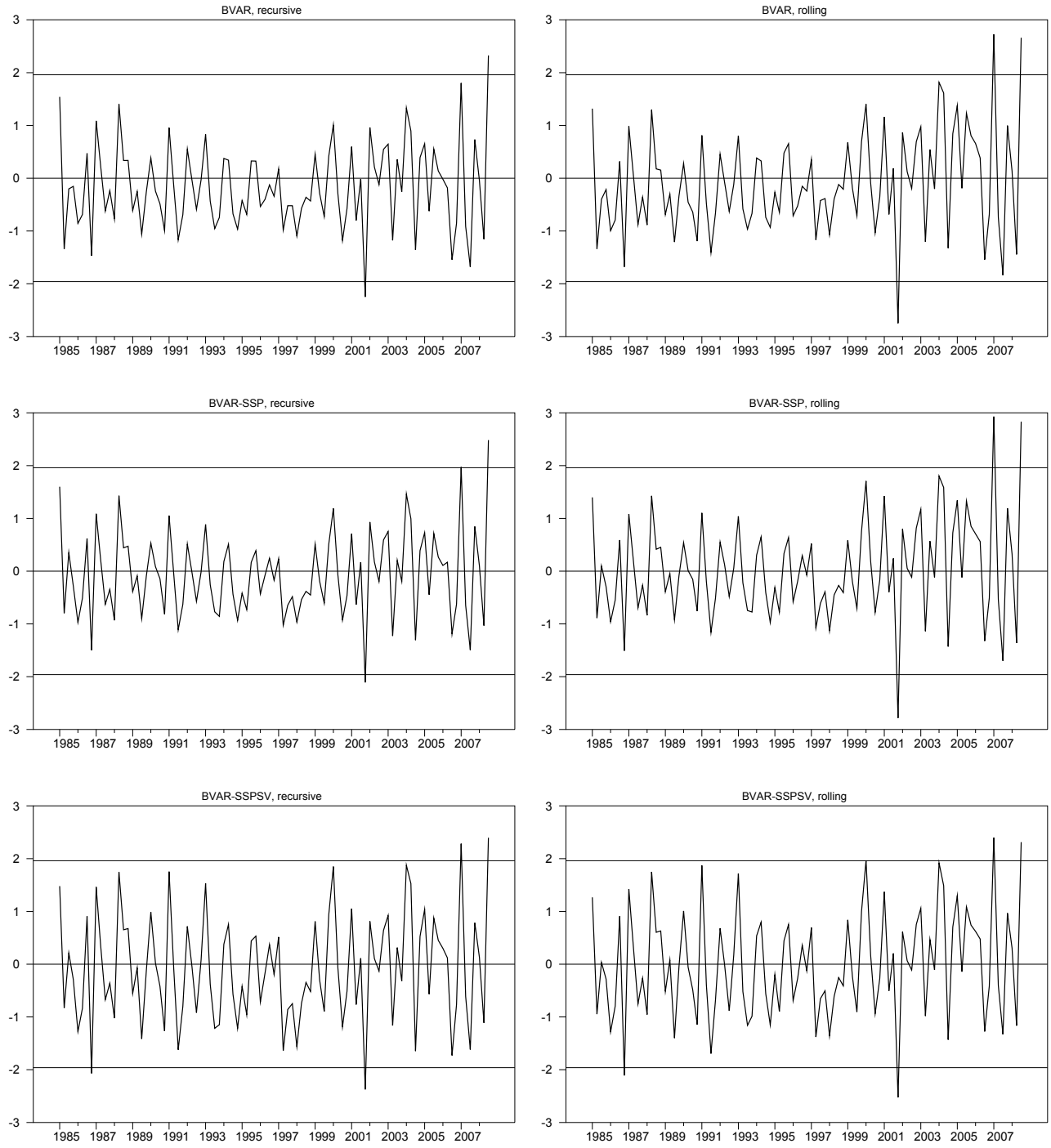
Notes: The normalized forecast errors shown are defined as $\Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of a one-step ahead forecast error (generated in real time) and Φ^{-1} is the inverse of the standard normal distribution function. The horizontal lines included in the charts represent 95 percent intervals for the normal distribution.

Figure 8. Normalized Forecast Errors for unemployment
horizon = 1Q



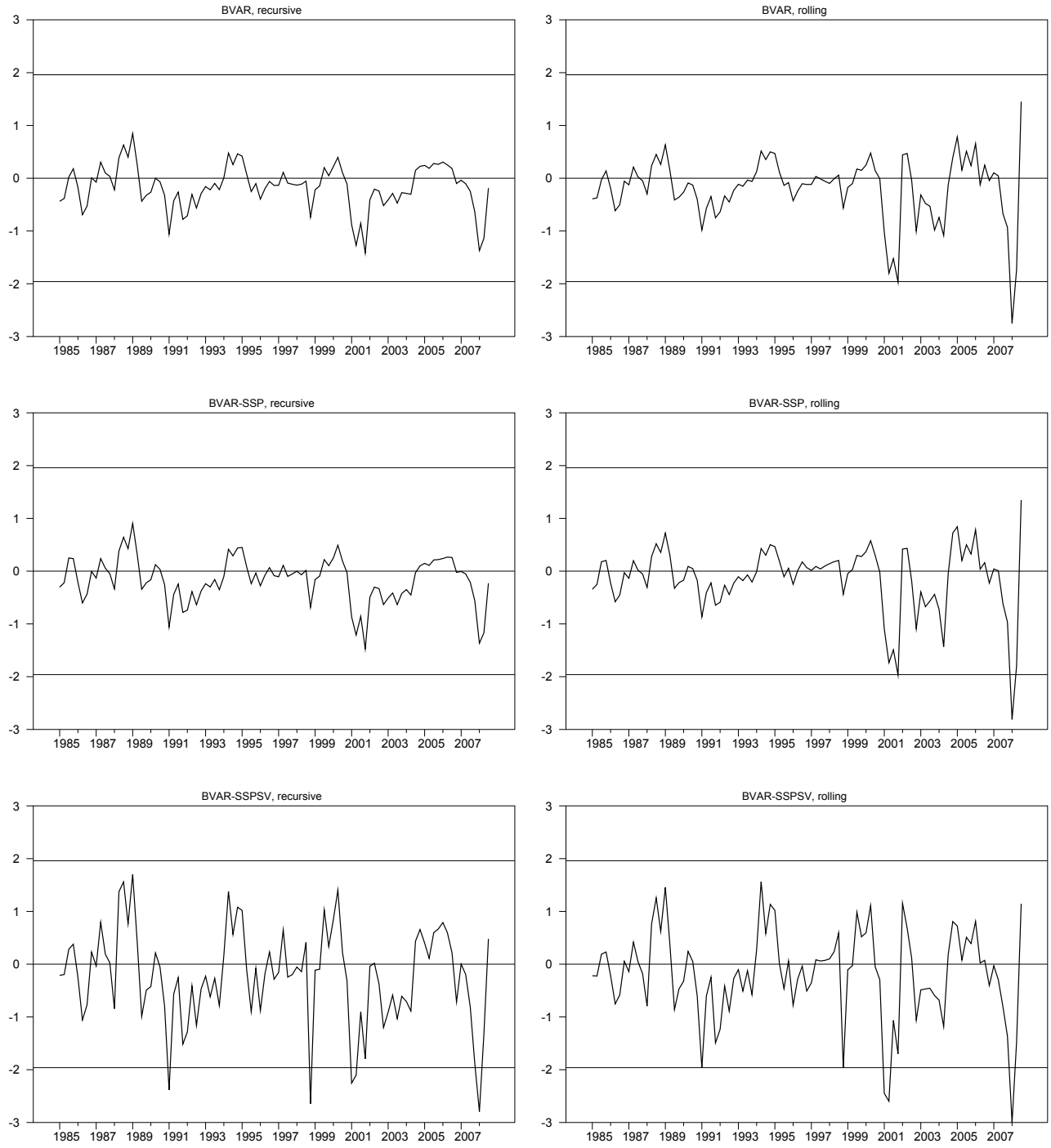
Notes: The normalized forecast errors shown are defined as $\Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of a one-step ahead forecast error (generated in real time) and Φ^{-1} is the inverse of the standard normal distribution function. The horizontal lines included in the charts represent 95 percent intervals for the normal distribution.

Figure 9. Normalized Forecast Errors for GDP inflation
horizon = 1Q



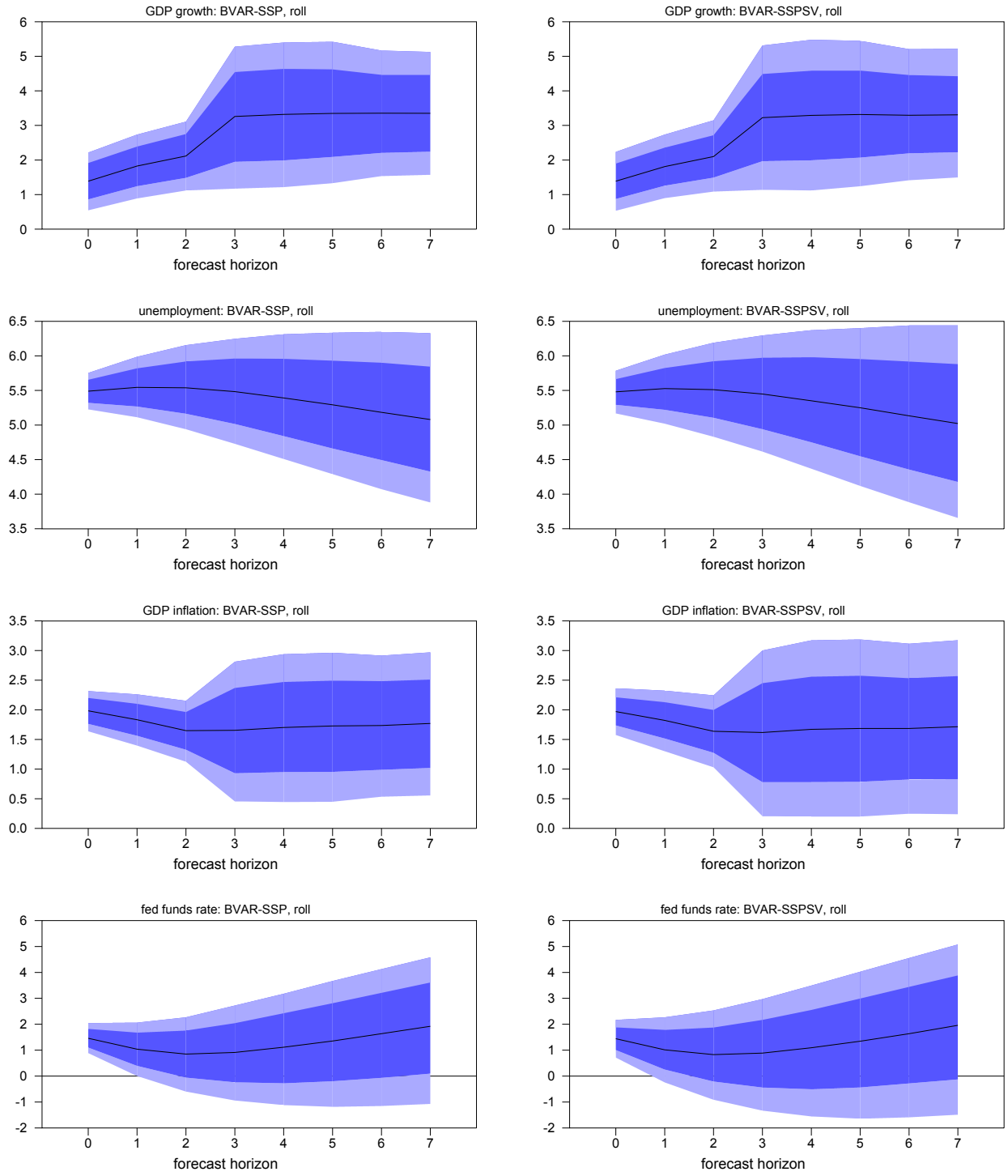
Notes: The normalized forecast errors shown are defined as $\Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of a one-step ahead forecast error (generated in real time) and Φ^{-1} is the inverse of the standard normal distribution function. The horizontal lines included in the charts represent 95 percent intervals for the normal distribution.

Figure 10. Normalized Forecast Errors for fed funds rate
horizon = 1Q



Notes: The normalized forecast errors shown are defined as $\Phi^{-1}(z_{t+1})$, where z_{t+1} denotes the PIT of a one-step ahead forecast error (generated in real time) and Φ^{-1} is the inverse of the standard normal distribution function. The horizontal lines included in the charts represent 95 percent intervals for the normal distribution.

Figure 11. Fan charts for forecasts made in 2008:Q3



Notes: The figure provides fan charts of forecasts (percentiles of the marginal density for each variable at each horizon) made in the middle of 2008:Q3, using the 2008:Q3 vintage of data, with a data sample ending in 2008:Q2. The GDP growth and inflation forecasts are reported as four-quarter averages. Period 0 on the horizontal axis of each chart refers to the forecast for 2008:Q3.