WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 10-19
EVALUATING REAL-TIME VAR FORECASTS WITH AN INFORMATIVE DEMOCRATIC PRIOR

Jonathan H. Wright
Johns Hopkins University and
Visiting Scholar, Federal Reserve Bank of Philadelphia

May 19, 2010

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Evaluating Real-Time VAR Forecasts with an Informative democratic Prior

Jonathan H. Wright*

First Draft: October 2009
This version: May 19, 2010

Abstract

This paper proposes Bayesian forecasting in a vector autoregression using a *democratic* prior. This prior is chosen to match the predictions of survey respondents. In particular, the unconditional mean for each series in the vector autoregression is centered around long-horizon survey forecasts. Heavy shrinkage toward the democratic prior is found to give good real-time predictions of a range of macroeconomic variables, as these survey projections are good at quickly capturing endpoint-shifts.

*Department of Economics, Johns Hopkins University, Baltimore MD 21218; wrightj@jhu.edu.
I am grateful to Todd Clark, Neil Ericsson, Jon Faust, Gary Koop, Lucrezia Reichlin and Frank Schorfheide for their helpful comments on earlier drafts. All errors are my sole responsibility. The views expressed in this paper are those of the author and do not necessarily reflect those of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. This paper is available free of charge at http://www.philadelphiafed.org/research-and-data/publications/working-papers/.

1 Introduction

Since the seminal paper of Sims (1980), vector autoregressions (VARs) have become widely used for forecasting in macroeconomics and finance. However, VARs can have a large number of parameters, and it was quickly recognized that forecast performance can accordingly be improved by adopting some kind of Bayesian shrinkage. This was the motivation of the Minnesota prior (Doan, Litterman and Sims (1984) and many papers since then).

This paper considers a Bayesian approach to forecasting in a VAR but with a quite different prior. The prior that I use is instead centered around the parameters that are implied by survey responses. Specifically, the prior is that each variable follows a univariate autoregression in which the unconditional mean is set equal to the most recent long-run (five-to-ten-year-ahead) survey forecast. The persistence of each autoregression may also be set to match the properties of the short-term survey predictions for that variable. I call this the "democratic" prior. I find that Bayesian forecasts that shrink heavily toward this prior have excellent real-time forecasting performance for a range of macroeconomic variables.

It seems natural to shrink toward survey forecasts. Survey or other judgmental forecasts have been found to have excellent properties relative to a large battery of econometric predictions, especially for nominal variables (see e.g. Ang, Bekaert and Wei (2007), Croushore (2008) and Faust and Wright (2009)). Indeed, one might just use the surveys as forecasts, but unfortunately surveys are conducted only infrequently, and cover only some variables and some horizons. Besides, having a model-based forecast is useful for understanding the mechanisms underlying the predictions that are being made. All this motivates forecasting from a time-series method, but using survey evidence to construct an informative prior.

The plan for the remainder of this paper is as follows. In section 2, I discuss some standard priors for VARs and introduce the proposed democratic prior. Section 3 reports the results from applying these forecasting methods with real-time data, including comparison with a range of alternative prediction techniques. Section 4 concludes.

2 Priors in Vector Autoregressions

Consider the VAR

$$y_t = k + A_1 y_{t-1} + A_2 y_{t-2} ... + A_p y_{t-p} + u_t \tag{1}$$

where y_t is an nx1 vector and u_t is i.i.d. $N(0, \Sigma)$.

It is standard to estimate (1) by OLS and then iterate forward to construct forecasts. Alternatively, Bayesian approaches can be used. A standard Bayesian approach would be to use the normal-diffuse prior, where the priors for k, $A = [A_1 \ A_2 \ ... \ A_p]$ and Σ are mutually independent with

$$p(k)\tilde{\ }N(0, \kappa I_n) \tag{2}$$

$$p(vec(A'))\tilde{\ }N(0_{np^2\text{x}1}, \Omega_A) \tag{3}$$

and

$$p(\Sigma) \propto |\Sigma|^{-(n+1)/2} \tag{4}$$

where κ is a large number, Ω_A is a diagonal matrix, the prior variance for the ijth element of A_k is $\frac{\lambda^2}{k^2}\frac{\sigma_i^2}{\sigma_j^2}$, λ is a hyperparameter that measures the overall tightness of the prior and σ_i^2 is the residual variance from fitting an AR(1) to y_{it}. The prior for k and A is thus a variant of the Minnesota prior of Doan, Litterman and Sims (1984) that was adopted by Banbura, Giannone and Reichlin (2010).

Write the VAR as a multivariate regression

$$Y = X\beta + U \tag{5}$$

where Y is a Txn matrix, X is of order Tx$(np^2 + 1)$, $B = [k \ A_1 \ A_2 \ ... \ A_p]'$ is of order $(np^2 + 1)$xn, and U is a Txn matrix of errors that are independent over time and $N(0,\Sigma)$. The prior for $vec(B')$ can be written as $N(0_{(p+np^2)\text{x}1}, \Omega_b)$, and the Gibbs sampler can then be used to take draws from the posterior of the parameters. Specifically, the posterior of $vec(B)$ conditional on Σ is

$$N((\Omega_b^{-1} + \Sigma^{-1} \otimes X'X)^{-1} vec(X'Y\Sigma^{-1}), (\Omega_b^{-1} + \Sigma^{-1} \otimes X'X)^{-1}) \tag{6}$$

while the posterior for Σ conditional on B is

$$IW((Y - XB)'(Y - XB), T) \tag{7}$$

where $IW(.,.)$ denotes the inverse-Wishart distribution (see Kadiyala and Karlsson (1997)). The Gibbs sampler builds up the posterior by iterating between equations (6) and (7).

The Minnesota prior has shown some improvements in forecasting relative to OLS estimation of the VAR. But one might search for other priors. For example, Ingram and Whiteman (1994) and del Negro and Schorfheide (2002) use dynamic stochastic general equilibrium models to construct priors. In this paper, I go in a different direction instead using the judgment of survey respondents to construct priors.

Blue Chip economic forecasting asks respondents to provide a long-term (five-to-ten-year ahead) forecast for a range of macroeconomic variables each March and October. This same survey asks respondents for near-term forecasts as well—quarterly predictions for each quarter in the current and next calendar year. For each variable, I take the long-run survey projection to be the survey-based estimate of mean of the variable, μ_i^*. I can also solve for the AR(1) slope coefficient for the ith variable that is implicit in the survey, ρ_i, from the optimization problem

$$\rho_i^* = \arg \min_{\rho_i \in [-0.99, 0.99]} \Sigma_{h=1}^{H} [(\hat{y}_{i,h} - \mu_i^*) - \rho_i(\hat{y}_{i,h-1} - \mu_i^*)]^2 \tag{8}$$

where $\hat{y}_{i,h}$ denotes the h-quarter ahead forecast for the ith variable. This is motivated by the fact that if the time series really were following an AR(1), then we would have $(\hat{y}_{i,h} - \mu_i^*) - \rho_i(\hat{y}_{i,h-1} - \mu_i^*)$. Let $\mu^* = (\mu_1^*, \mu_2^*, ...\mu_n^*)'$ and $\rho^* = (\rho_1^*, \rho_2^*, ...\rho_n^*)'$. I use Blue Chip forecasts rather than the forecasts from the Survey of Professional Forecasters (SPF) because the SPF has much less information on long-run expectations.[1]

The Minnesota prior is designed not to be informative about the intercept. Nevertheless, I would argue that our strongest prior beliefs from surveys are precisely about the mean of macroeconomic time series. For implementing a prior on the

[1]Blue Chip has long-term forecasts for about ten variables twice a year. The SPF did not begin asking any questions about long-term expectations until 1991Q4. Then, it asked questions about CPI inflation every quarter and asked about four other variables in the first quarter of each year only. Recently, the SPF has expanded its long-term questions further, but the sample period is too short for a forecast evaluation exercise of the kind considered in this paper.

mean based on this survey information, it will be convenient to adopt a different parameterization of the VAR:

$$(y_t - \mu) = A_1(y_{t-1} - \mu) + A_2(y_{t-2} - \mu).... + A_p(y_{t-p} - \mu) + u_t \tag{9}$$

Clearly equations (1) and (9) are equivalent, with $\mu = (I - A_1 - A_2.... - A_p)^{-1}k$. However, equation (9) is a more convenient form for representing prior beliefs about the mean of the series. Villani (2009) developed the technology for a Bayesian analysis of equation (9).

Thus, following Villani (2009), consider the prior for the VAR in which μ, A and Σ are independent and:

$$p(vec(A)')\tilde{}N(\theta_A, \Omega_A) \tag{10}$$

$$p(\mu)\tilde{}N(\theta_\mu, \Omega_\mu) \tag{11}$$

and

$$p(\Sigma)\tilde{}|\Sigma|^{-(n+1)/2} \tag{12}$$

In equation (10), the prior mean of A_1 is set to $diag(\rho_1, \rho_2, ...\rho_n)$, while the remaining elements of θ_A are equal to zero, and Ω_A is as before a diagonal matrix with the prior variance for the ijth element of A_k being set to $\frac{\lambda^2}{k^2}\frac{\sigma_i^2}{\sigma_j^2}$. The prior variance for the mean is $\Omega_\mu = \lambda_0 I_n$. The posterior now involves a Gibbs sampler iterating between three steps provided by Villani.[2]

[2]The prior for the steady state should be reasonably informative (small λ_0), because this gives good out-of-sample forecasting performance, as I shall show later. Moreover, Villani (2009) points out that if the VAR is nearly nonstationary *and* the steady-state prior is very uninformative, then the Gibbs sampler will have convergence problems.

The *democratic* prior that I propose in this paper is of the form of equations (10)-(12), letting $\theta_\mu = \mu^*$, the long-run survey means. All that remains is to specify the prior means of the autoregressive slope coefficients $(\rho_1, \rho_2, ...\rho_n)$. For this, I consider three alternative approaches, giving three variants on the democratic prior:

Prior D1 sets $\rho_1 = \rho_2.... = \rho_n = 0$, as in Banbura, Giannone and Reichlin (2010).

Prior D2 sets $\rho_i = 0$ for each real variable, but $\rho_i = 0.8$ for each nominal variable.

Prior D3 sets $\rho_i = \rho_i^*$, the slope coefficient obtained from the optimization problem in (8).

The two hyperparameters of the model are λ and λ_0, which determine the overall tightness of the prior for the slope coefficients and mean, respectively.

Research that considers the role of judgment in forecasting faces the problem that judgment typically gives us information about future data points, whereas what we would really like is non-model information about the parameter value (see, for example Manganelli (2009)). An advantage of the parameterization in (9) combined with the long-horizon Blue Chip forecast is that they are one and the same thing—the long-term Blue Chip prediction is a survey-based estimate of the parameter μ. It is unfortunately not quite so easy for the autoregressive slope coefficients, but the prior information on those turns out to be far less important anyway.

Of course, Blue Chip survey respondents had access to the same time series as the vector autoregression in making their forecasts. To the extent that their survey forecasts are influenced by these data, the Bayesian VAR with a democratic prior is effectively using the data as a prior, which of course violates the principles of Bayesian statistical inference. However, the information underlying the survey is much richer than the series in the VAR. Besides, I am simply viewing the democratic prior as a pragmatic forecasting device, that should be assessed purely on the basis of its predictive performance. I turn to evaluating this in the next section.

3 Implementation and Real-Time Out-of-Sample Forecasting

The VAR considered in this paper is a medium-size system consisting of ten quarterly macroeconomic variables: real GDP growth, real GDP deflator inflation, CPI inflation, industrial production growth, growth in nonresidential fixed investment, growth in real personal consumption expenditures, ten-year yields, three-month Treasury bill yields, housing starts and the unemployment rate. Yields, housing starts and unemployment rates are in levels; the other variables are all annualized growth rates.[3] For each quarterly vintage of data from the Federal Reserve Bank of Philadelphia's real-time dataset from 1984Q2 to 2009Q1, I estimated a VAR(4) in these variables using data from 1960Q1 onward as observed in that data vintage. The VAR was estimated in the following ways:

(i) By OLS,

(ii) Using the Minnesota prior in equations (2)-(4), and

(iii) Using the democratic prior proposed in this paper (variants D1, D2, and D3), employing in each case the most recent Blue Chip survey data.[4]

---

[3] Specifically 400 times log first-differences.

[4] That is, for forecasts made in the second and third quarters, the Blue Chip survey used is from the previous March, while for forecasts made in the first or fourth quarters, the October survey is

The forecasting exercise is out-of-sample and fully real-time. The data used for forecasting in quarter t are the data as observed in the middle of that quarter and consist of the observations from 1960Q1 up to and including quarter $t-1$ for all ten variables. Throughout this paper, I consider the forecasts of six variables: real GDP growth, GDP deflator and CPI inflation, industrial production growth, three-month yields and the unemployment rate, at horizons ranging from 0 to 13 quarters.[5] Real-time forecasting exercises require some definition of the "actual" data—the definition used in this paper is the data that are observed in the middle of the second quarter after the quarter to which they refer. For example, the "actual" values for 2009Q1 are the values observed in August of 2009. Table 1 shows the root-mean-square error (RMSE) of the OLS VAR forecasts.

The Bayesian forecasts depend on hyperparameters: λ in the case of the Minnesota prior, and both λ and λ_0 in the case of the democratic priors. Figure 1 plots the RMSE of the forecasts using the Minnesota prior relative to the RMSE of the OLS VAR forecast against λ. Figures 2-4 plot the relative RMSE of the forecasts using democratic priors D1, D2 and D3, respectively against λ_0. In all cases the horizons considered are current-quarter and one-, four- and eight-quarters hence. In implementing the democratic prior, I fix λ at 0.1.

Turning first to Figure 1, the Minnesota prior can bring gains in forecast precision, especially with small values of λ. Relative to the OLS benchmark, the RMSE can

---

used. Each Blue Chip survey gives a long-run mean and an implied autoregressive parameter for each variable.

[5] Results for the other series are omitted to conserve space.

be reduced by as much as 20 percent. But, as shown in Figures 2-4, the democratic prior gives substantially larger reductions in RMSE.[6] The forecasting improvements using the democratic prior are most notable for inflation, and especially at long horizons. At the eight-quarter horizon, using the democratic prior yields more than a 40 percent reduction in RMSE. For inflation, the improvement is largest when λ_0 is small, meaning that heavy shrinkage is optimal. The democratic prior works well for forecasting other variables too. For forecasting real GDP growth, using the democratic prior results in a 5-10 percent reduction in RMSE (relative to OLS) at horizons up to and including four quarters.

All three variants of the democratic prior beat OLS in forecast accuracy for nearly all variables and forecast horizons—and in the few cases in which this is not true, it is for all practical purposes a tie. The results using democratic priors D1, D2 and D3 are very similar to each other. This means that shrinkage toward the long-run survey forecast is key—virtually all of the improvement in the democratic prior comes from simply getting the end-point right. However, D2 seems to give slightly more accurate forecasts than D1, and D3 does better again. So although the device for estimating implied persistence in equation (8) is rather *ad hoc*, it still seems to help a little with forecasting.

Figure 5 shows the Diebold-Mariano t-statistics testing the hypothesis that the RMSE of forecasts estimating the VAR by OLS and democratic prior D3 are

---

[6] In democratic prior D3, the autoregressive parameter estimates implied by the surveys (the solutions to equation (8)) are volatile, but typically lie between 0.6 and 0.99 for all the variables being forecast in this paper. The autoregressive parameter estimates tend to be a little larger for inflation and other nominal indicators than for real indicators.

equal (results comparing OLS with democratic priors D1 and D2 are similar, and not shown). These t-statistics are formed as described by Diebold and Mariano (1995) and are based on Newey-West standard errors with a lag truncation parameter equal to the forecast horizon. The t-statistics are again plotted against λ_0. The t-statistics are significant[7] at least at the 10 percent level at all horizons for both inflation measures (and generally at the 5 percent level). Significant values are also obtained for real GDP and industrial production growth at some horizons.

3.1 Bias in the forecasts and shifting endpoints

Estimating the VAR by OLS, or using the Minnesota prior, the forecast at distant horizons converges to the sample mean of the time series over the estimation period (the Minnesota prior uses a diffuse prior for the intercept). Meanwhile, the democratic prior gives a forecast that converges to a mixture of the sample average and the long-run survey forecast. If there are occasional shifts in the mean of the time series, along the lines of the "shifting endpoints" considered by Kozicki and Tinsley (2001)[8], or the intercept shifts considered by Clements and Hendry (1998, 1999), and the survey respondents are aware of these shifts in real time, then this will make the democratic prior perform better particularly at longer horizons. In effect, the democratic prior is a rapidly adjusting and robust way of allowing for intercept shifts.

As an illustration, Figure 6 shows the time series of eight-quarter-ahead forecasts

[7] Here and throughout this paper, I am considering tests against a two-sided alternative, comparing the t statistics with standard normal critical values. Normal critical values are appropriate asymptotically if the Bayesian and OLS forecasts are thought of as non-nested.

[8] In particular, Kozicki and Tinsley (2001) find that it is easier to rationalize the properties of Treasury yields in a model in which the Fed's implicit inflation target is subject to permanent shocks.

of inflation (GDP deflator) from OLS, the Minnesota prior, and democratic prior D3. The actual value of inflation eight quarters later is also plotted. The OLS estimates of the VAR are stationary and hence by construction give forecasts of inflation that were rising up toward the sample mean of inflation since 1960. Of course, these predictions turned out to be consistently too high. The Minnesota prior consistently overpredicted inflation, for the same reason. The democratic prior did substantially better, because long-term survey forecasts of inflation declined sharply over the 1980s and early 1990s.

The point is shown more generally in Table 2, which shows the bias of the OLS forecasts and the forecasts using the Minnesota and democratic priors (D1, D2 and D3) with $\lambda = 0.1$ and $\lambda_0 = 0.05$. The forecasts for inflation and real activity are generally upwardly biased over this period because of the disinflation of the 1980s and 1990s and the productivity slowdown in the 1970s. This is true for the democratic prior as well. But, because the survey respondents learned about the disinflation and productivity slowdown reasonably quickly, all variants of the democratic prior have substantially smaller bias than OLS or the Minnesota prior, especially at longer horizons.[9]

3.2 Setting the prior parameters in a training sample

The Minnesota and democratic priors depend on nuisance parameters: λ and λ_0. Viewing Bayesian forecasting as a pragmatic shrinkage device a natural approach is

[9] This naturally motivates thinking of non-stationary models. Running a VAR in which inflation enters in first differences does greatly mitigate the bias of the inflation forecasts, but it increases the variance and the RMSE. Other non-stationary models are discussed in subsection 3.4 below.

to select the values of these nuisance parameters to minimize RMSE over a training sample and then to hold them fixed at these values over all subsequent forecasts.

To implement this, I picked the values of the hyperparameters λ and λ_0 to minimize out-of-sample RMSE in the data as observed at the end of 1990 (a fairly short sample). And then, I used these values of λ and λ_0 for forecasting starting in 1991Q1, holding the hyperparameters fixed at these chosen values for the remainder of the sample period. Table 3 shows the resulting out-of-sample RMSE using the Minnesota prior and democratic priors D1, D2 and D3, relative in all cases to the RMSE from simply using OLS.

In some cases, notably long-horizon forecasting of inflation, the democratic prior gives gains over OLS that are both economically and statistically significant. In other cases, it is about a tie, but in no case does the democratic prior do much worse. The gains are roughly comparable for all three versions of the democratic prior, but overall D3 does a little better than D2, which in turn has a small edge over D1. The Minnesota prior also gives consistent, but more modest, gains. All this is in line with the results from Figures 1-4, but avoids any dependence on somewhat arbitrarily chosen nuisance parameters.

3.3 Comparing the democratic prior with surveys

The generally strong performance of the VAR with the democratic prior in turn begs the question of how good the forecasting performance would be if we simply discarded the VAR and instead used the surveys—at least for those horizons for which a survey forecast is available. Table 4 accordingly compares the RMSE of the VAR forecasts obtained from OLS, the VAR forecasts using the Minnesota prior (with λ =0.1), the VAR forecasts using the three versions of democratic prior (with $\lambda = 0.1$ and $\lambda_0 = 0.05$), and the Blue Chip survey forecasts. These are the short-term Blue Chip survey forecasts for the next few quarters—not the five-to-ten-year ahead survey projections on which the democratic prior is based.

With any comparison between survey and time-series forecasts, thorny issues of timing arise. If we compare a VAR forecast based on data up to and including quarter $t-1$ with a survey taken during quarter t, then the survey respondents had access to some information from quarter t, giving the survey an artificial timing advantage. On the other hand, if we use a survey taken during quarter $t-1$, then the survey respondents could have known only part of the data for quarter $t-1$, putting them at an unfair timing disadvantage. There is no way to structure the comparison so that the survey and VAR forecasts are based on identical information sets.[10] In Table 4, I adopt the convention of using the survey forecasts from the last month of quarter $t-1$, and compare these forecasts with VAR predictions that condition on all information up to and including quarter $t-1$. The forecasts are for quarters t, $t+1$, $t+2$, and $t+3$: Quarterly survey forecasts are not always available at longer horizons. This timing convention means that the surveys are being put at something of a timing *disadvantage*.

Notwithstanding this timing disadvantage, the surveys do very well in the comparison in Table 4. The RMSE from the raw surveys is about the same as that

[10] Early in quarter t, survey respondents will have some data for quarter t, but not all the data for quarter $t-1$, and so it is not clear whether the survey is at a timing advantage or disadvantage. But the information sets in the VAR and survey forecasts are still not identical.

from the VAR with democratic prior D3—in some cases slightly higher, in other cases slightly lower. All in all, using survey data seems an excellent approach to forecasting and is very hard for econometric models to beat convincingly (as found by Ang, Bekaert and Wei (2007) and Croushore (2008)). But VARs are a central tool in empirical macroeconomics used for purposes other than forecasting: including impulse response analysis, variance decompositions, or consideration of counterfactual scenarios, none of which can be undertaken from the raw survey predictions alone. Also, the VAR allows econometricians to make forecasts at horizons not considered in the survey, for variables not predicted in the survey and to make projections at times when the survey is not being taken. Nonetheless, some researchers are reluctant to use VARs because their out-of-sample forecasting performance can be quite poor. The results in Table 4 indicate that this problem is solved by the use of the democratic prior: The democratic prior gets the VAR forecasts to be able to roughly match a tough benchmark.

3.4 Comparison with time-varying parameter methods

The democratic prior seems to work well because of its ability to adjust to intercept shifts. There are of course many more-standard econometric approaches to forecasting in the presence of possible structural changes. In this subsection, I compare forecasts using the democratic prior with some of these existing alternatives. Throughout this subsection, I am considering variant D3 of the democratic prior, with $\lambda = 0.1$ and $\lambda_0 = 0.05$.

I first consider forecasting inflation using the unobserved components stochastic volatility (UCSV) model proposed by Stock and Watson (2007). The model is a univariate specification that inflation is $\pi_t = \pi_t^T + \pi_t^P$ where $\pi_t^P = \pi_{t-1}^P + \xi_t$ and π_t^T and ξ_t are martingale difference sequences with stochastic volatility. The long-run component of inflation is π_t^P, and it is time-varying. The model can be estimated by Markov Chain Monte Carlo methods. Table 5 shows the out-of-sample RMSE of UCSV forecasts of inflation (CPI and GDP deflator) and also reports the out-of-sample RMSE of inflation forecasts using the ten-variable VAR with the democratic prior. The UCSV model is a good benchmark, and gives more accurate long-horizon inflation forecasts than are obtained from OLS estimation of the VAR (see Table 1). Nonetheless, comparing the UCSV model and the VAR with the democratic prior, the latter gives reductions in RMSE for forecasting inflation that are both economically and statistically significant.

VARs with drifting parameters have become popular recently (Cogley and Sargent (2005), Primiceri (2005) and Cogley, Primiceri and Sargent (2010)). For computational reasons, they are used only with relatively small systems. For example, Primiceri (2005) considered a VAR with time-varying parameters and stochastic volatility and applied it to a system with three variables: inflation, the unemployment rate and the federal funds rate. I applied Primiceri's VAR to forecasting GDP deflator inflation, the unemployment rate and three-month Treasury bill yields.[11] Table 6 compares the out-of-sample RMSE of these time-varying VAR forecasts with the forecasts from the ten-variable VAR using the democratic prior. The two forecasts are roughly comparable in terms of forecast accuracy, with the democratic prior VAR

[11] The priors are exactly as in Primiceri (2005). The lag order is 2, again following Primiceri.

having a slight edge in most (but not all) cases.

Finally, another potential approach for forecasting in the presence of parameter instability is to use OLS estimation of a VAR but to do so in a rolling sample. This is a simple and yet widely-used method to account for the possibility of structural breaks without introducing any element of judgment from surveys. To assess how well a rolling-coefficients VAR works, I computed the out-of-sample RMSE of the VAR forecasts using a 40-quarter rolling window. The results are shown in Table 7, along with the out-of-sample RMSE of corresponding forecasts using the democratic prior[12]. As can be seen in Table 7, using the rolling VAR is generally less accurate, and significantly so in some cases.[13]

4 Conclusion

Surveys give good predictions of many macroeconomic variables, perhaps in part because they are better able to adapt to low-frequency structural breaks than any statistical time series model. This seems particularly true for inflation forecasts. Nonetheless, time series forecasts are useful for a number of reasons, including the fact that they apply to variables and horizons that are not included in the survey, and can be worked out at any time. Besides, having a model-based forecast is useful for understanding the mechanisms underlying the predictions that are being made.

[12] In this exercise, both VARs use ten variables. The rolling VAR is of order 1, because it would run out of degrees of freedom if 4 lags were included.

[13] An alternative to the rolling VAR is to consider a VAR in which there is an intercept shift forty quarters before the end of the sample (but no break in the other parameters). This also generally gives less accurate out-of-sample forecasts than the VAR using the democratic prior.

Models can be used for purposes such as impulse response analysis, or examining counterfactual scenarios, none of which can be obtained from raw surveys alone. It thus seems natural to estimate a VAR, but to do so with an informative Bayesian prior that shrinks toward the values implied by surveys.

In this paper I have proposed a concrete way of implementing such a Bayesian VAR. In a real-time forecasting exercise, I have found that it often outperforms both OLS estimation of the VAR and Bayesian VAR estimation using the Minnesota prior. The improvements are most consistent for inflation forecasting at longer horizons, and appear to owe mainly to the ability of the surveys to capture shifting end-points.

References

Ang, A., G. Bekaert and M. Wei (2007): Do Macro Variables, Asset Markets or Surveys Forecast Inflation Better?, *Journal of Monetary Economics*, 54, pp.1163-1212.

Banbura, M., D. Giannone and L. Reichlin (2010): Large Bayesian VARs, *Journal of Applied Econometrics*, 25, pp.71-92.

Clark, T.E. (2009): Real-Time Density Forecasts from VARs with Stochastic Volatility, Federal Reserve Bank of Kansas City Working Paper 09-08.

Clements, M.P. and D.F. Hendry (1998): "Forecasting Economic Time Series," Cambridge University Press, Cambridge, England.

Clements, M.P. and D.F. Hendry (1999): "Forecasting Nonstationary Economic Time Series," MIT Press, Cambridge, Massachusetts.

Cogley, T., G.E. Primiceri and T.J. Sargent (2010): Inflation-Gap Persistence in the U.S., *American Economic Journal: Macroeconomics*, 2, pp.43-69.

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

Croushore, D. (2008): An Evaluation of Inflation Forecasts from Surveys Using Real-Time Data. Working paper.

Del Negro, M. and F. Schorfheide (2002): Priors from General Equilibrium Models for VARs, *International Economic Review*, 45, pp.643-673.

Diebold, F.X. and R.S. Mariano (1995): Comparing Predictive Accuracy, *Journal of Business and Economic Statistics*, 13, pp.253-263.

Doan, T., R. Litterman and C.A. Sims (1984): Forecasting and Conditional Projection Using Realistic Prior Distributions, *Econometric Reviews*, 3, pp.1-100.

Faust, J. and J.H. Wright (2009): Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset, *Journal of Business and Economic Statistics*, 27, pp.468-479.

Ingram, B. and C. Whiteman (1994): Supplanting the Minnesota prior—Forecasting Macroeconomic Time Series Using Real Business Cycle Model Priors, *Journal of Monetary Economics*, 34, pp.497-510.

Kadiyala, K.R. and S. Karlsson (1997): Numerical Methods for Estimation and Inference in Bayesian VAR models, *Journal of Applied Econometrics*, 12, pp.99-132.

Kozicki, S. and P.A. Tinsley (2001): Shifting Endpoints in the Term Structure of Interest Rates, *Journal of Monetary Economics*, 47, pp.613-652.

Manganelli, S. (2009): Forecasting with Judgment, *Journal of Business and Economic Statistics*, 27, pp.553-563.

Primiceri, G.E. (2005): Time Varying Structural Vector Autoregressions and Monetary Policy, *Review of Economic Studies*, 72, pp.821-852.

Sims, C.A. (1980): Macroeconomics and Reality, *Econometrica*, 48, pp.1-48.

Stock, J.H. and M.W. Watson (2007): Why has U.S. Inflation Become Harder to Forecast?, *Journal of Money, Credit and Banking*, 39, pp.3-33.

West, K.D. (2006): Forecast Evaluation, in "Handbook of Economic Forecasting," Vol. 1, G. Elliott, C.W.J. Granger and A. Timmerman (eds.), Elsevier, Amsterdam, The Netherlands.

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

Zellner, A. (1971): An Introduction to Bayesian Inference in Econometrics, Wiley, New York.

Table 1: RMSE of VAR forecasts Using OLS

Horizon (Quarters)	h=0	h=1	h=4	h=8
Real GDP Growth	2.44	2.63	2.65	2.41
GDP Deflator Inflation	1.22	1.28	1.74	2.44
CPI Inflation	2.22	2.56	2.84	3.31
IP Growth	5.08	5.72	5.77	5.02
TBill Yields	0.88	1.22	1.96	2.56
Unemployment Rate	0.32	0.49	0.81	0.97

Notes: This table reports the out-of-sample root-mean-square error (RMSE) of forecasts from real-time OLS estimation of the ten-variable VAR, as described in the text. The sample period is 1984Q2-2009Q1. The units of real GDP growth, GDP deflator inflation, CPI inflation and IP growth are annualized quarter-over-quarter percentage changes (more precisely, 400 times log first differences). The units of yields and the unemployment rate are percentage points.

Table 2: Bias from Alternative VAR Estimates

Horizon (Quarters)		h=0	h=1	h=4	h=8
Real GDP Growth	OLS	0.86	1.17	0.77	0.33
	Minnesota	0.97	1.31	1.15	0.73
	Democratic: D1	0.57	0.77	0.66	0.52
	Democratic: D2	0.57	0.75	0.69	0.58
	Democratic: D3	0.39	0.61	0.64	0.53
GDP Deflator Inflation	OLS	0.34	0.53	1.21	1.91
	Minnesota	0.39	0.62	1.07	1.64
	Democratic: D1	0.26	0.38	0.59	0.77
	Democratic: D2	0.08	0.23	0.42	0.56
	Democratic: D3	0.10	0.24	0.44	0.58
CPI Inflation	OLS	0.23	0.41	1.12	1.94
	Minnesota	-0.01	0.23	0.76	1.46
	Democratic: D1	-0.11	0.07	0.25	0.46
	Democratic: D2	-0.09	0.03	0.09	0.26
	Democratic: D3	-0.08	-0.01	0.08	0.25
IP Growth	OLS	1.98	2.38	1.98	1.09
	Minnesota	2.13	2.78	2.67	1.76
	Democratic: D1	1.58	1.94	1.82	1.51
	Democratic: D2	1.49	1.89	1.87	1.62
	Democratic: D3	1.18	1.70	1.86	1.61
TBill Yields	OLS	0.15	0.27	0.78	1.38
	Minnesota	0.17	0.36	0.86	1.34
	Democratic: D1	0.28	0.46	0.80	1.07
	Democratic: D2	0.20	0.35	0.68	0.92
	Democratic: D3	0.15	0.29	0.62	0.87
Unemployment Rate	OLS	-0.04	-0.09	-0.20	-0.11
	Minnesota	-0.03	-0.08	-0.23	-0.23
	Democratic: D1	-0.02	-0.06	-0.12	-0.09
	Democratic: D2	-0.02	-0.06	-0.13	-0.13
	Democratic: D3	-0.02	-0.06	-0.12	-0.13

Notes: This table shows the out-of-sample mean forecast error (actual-predicted) for the different variables, horizons and forecasting methods. The sample period is 1984Q2-2009Q1 and the VAR contains ten variables, as described in the text. The Bayesian methods set λ=0.1 and λ_0=0.05.

Table 3: RMSE of Bayesian VAR Forecasts with Hyperparameters Set in a Training Sample (Relative to OLS Estimation of the VAR)

Horizon (Quarters)		h=0	h=1	h=4	h=8
Real GDP Growth	Minnesota	0.97	0.96	0.91**	0.99
	Democratic: D1	0.94	0.94	0.88*	0.99
	Democratic: D2	0.91	0.92	0.88*	0.99
	Democratic: D3	0.86**	0.90*	0.91**	0.99
GDP Deflator Inflation	Minnesota	0.93*	1.03	0.92*	0.99
	Democratic: D1	0.89**	0.92	0.68***	0.57***
	Democratic: D2	0.95	1.00	0.71***	0.62***
	Democratic: D3	0.91	0.95	0.68***	0.59***
CPI Inflation	Minnesota	0.95	0.92***	0.88***	0.95
	Democratic: D1	0.94	0.88***	0.79***	0.75**
	Democratic: D2	1.01	0.93*	0.81***	0.76**
	Democratic: D3	0.95	0.89***	0.81***	0.77**
IP Growth	Minnesota	0.98	0.98	0.93*	1.02
	Democratic: D1	0.91	0.94	0.91**	1.04
	Democratic: D2	0.89*	0.92	0.91**	1.04
	Democratic: D3	0.82***	0.89**	0.92**	1.03
TBill Yields	Minnesota	0.97	1.00	1.04	1.01
	Democratic: D1	0.94	0.96	0.96	0.89*
	Democratic: D2	0.82**	0.88	0.95	0.90
	Democratic: D3	0.76***	0.84	0.96	0.92
Unemployment Rate	Minnesota	0.98	0.98	1.00	0.98
	Democratic: D1	1.03	1.02	0.99	1.01
	Democratic: D2	1.00	0.99	0.98	1.02
	Democratic: D3	0.96	0.97	0.98	1.01

Notes: This table reports the out-of-sample root-mean-square error (RMSE) of forecasts from real-time Bayesian estimation of the ten-variable VAR described in the text over the period from 1991Q1-2009Q1. The values of the hyperparameters λ and λ_0 are set at those values that minimize RMSE over the training sample that consists of data starting 1984Q2 that was available in 1991Q4. In all cases, RMSE values are relative to those from OLS estimation of the VAR. In each case, a Diebold-Mariano test of the hypothesis that the population relative RMSE is equal to one was conducted. One, two, and three asterisks denote cases in which this test rejected the null hypothesis of equal forecast accuracy at the 10, 5, and 1 percent significance levels, respectively.

Table 4: Comparison of RMSE

Variable	Horizon	VAR					Surveys
		OLS	Minn	D1	D2	D3	
Real GDP	0	2.44	2.32	2.21	2.20	2.09	1.90
	1	2.63	2.53	2.42	2.40	2.34	2.12
	2	2.55	2.51	2.42	2.41	2.42	2.25
	3	2.64	2.42	2.35	2.36	2.37	2.28
GDP Inflation	0	1.22	1.20**	1.09	1.15	1.11	0.96
	1	1.28	1.28***	1.07	1.11	1.07	1.06
	2	1.39	1.37***	1.07	1.07	1.06	1.10
	3	1.51	1.51***	1.14	1.12	1.11	1.23
CPI Inflation	0	2.22	2.04	2.03	2.16	2.05	1.85
	1	2.56	2.17	2.11	2.22	2.12	2.03
	2	2.60	2.16	2.06	2.10	2.08	2.05
	3	2.69	2.32	2.15	2.25	2.20	2.11
IP Growth	0	5.08	4.66	4.36	4.27	3.98	4.03
	1	5.72	5.36	5.04	4.97	4.84	4.69
	2	5.43	5.37	5.11	5.08	5.07	4.88
	3	5.54	5.29	5.05	5.05	5.08	4.89
TBill Yields	0	0.88	0.71	0.74	0.65	0.62	0.66
	1	1.22	1.04	1.07	0.99	0.95	1.02
	2	1.40	1.28	1.31	1.26	1.23	1.36
	3	1.70	1.52	1.54	1.51*	1.48*	1.68
Unemployment	0	0.32	0.31**	0.32**	0.31***	0.30***	0.36
	1	0.49	0.47*	0.48*	0.47*	0.46*	0.50
	2	0.62	0.61*	0.62	0.61*	0.61*	0.63
	3	0.73	0.70*	0.73	0.71*	0.72*	0.74

Notes: This table reports the out-of-sample root-mean-square error (RMSE) of forecasts using (i) OLS estimation of the ten-variable VAR described in the text over the period from 1984Q2-2009Q1, (ii) estimation of the same VAR using the Minnesota prior with λ=0.1, (iii) estimation of the VAR using the democratic priors D1-D3 with λ=0.1 and λ_0=0.05, and (iv) the raw Blue Chip surveys. In these results, the raw Blue Chip surveys are given an artificial timing *disadvantage*, as described in the text. For each of the five VAR-based forecasts, a Diebold-Mariano test of the hypothesis that the population RMSE of that forecast is equal to the population RMSE of the corresponding survey forecast was conducted. One, two, and three asterisks denote cases in which this test rejected the null hypothesis of equal forecast accuracy at the 10, 5, and 1 percent significance levels, respectively.

Table 5: RMSE from UCSV Model and VAR with democratic Prior

Horizon (Quarters)		h=0	h=1	h=4	h=8
GDP Deflator Inflation	VAR: Dem Prior	1.11***	1.07***	1.13***	1.33**
	UCSV	1.77	1.79	1.79	1.93
CPI Inflation	VAR: Dem Prior	2.05***	2.12***	2.13***	2.19**
	UCSV	2.51	2.63	2.67	2.78

Notes: This table reports the out-of-sample forecast root-mean-square error (RMSE) from estimation of the ten-variable VAR using the democratic prior D3 with λ=0.1 and λ_0=0.05, and from estimation of the univariate UCSV model of Stock and Watson (2007). The sample period is 1984Q2-2009Q1. For each horizon and inflation measure, a Diebold-Mariano test of the hypothesis that the two forecasts have equal population RMSE was conducted. One, two, and three asterisks denote cases in which this test rejected the null hypothesis of equal forecast accuracy at the 10, 5, and 1 percent significance levels, respectively.

Table 6: RMSE from TVP-VAR Model and VAR with democratic Prior

Horizon (Quarters)		h=0	h=1	h=4	h=8
GDP Deflator Inflation	VAR: Dem Prior	1.11	1.07	1.13	1.33
	TVP-VAR	1.04	1.07	1.19	1.38
Unemployment Rate	VAR: Dem Prior	0.30	0.46*	0.80	1.00*
	TVP-VAR	0.32	0.50	0.85	1.14
TBill Yields	VAR: Dem Prior	0.62*	0.95	1.71**	2.29*
	TVP-VAR	0.55	0.93	1.86	2.53

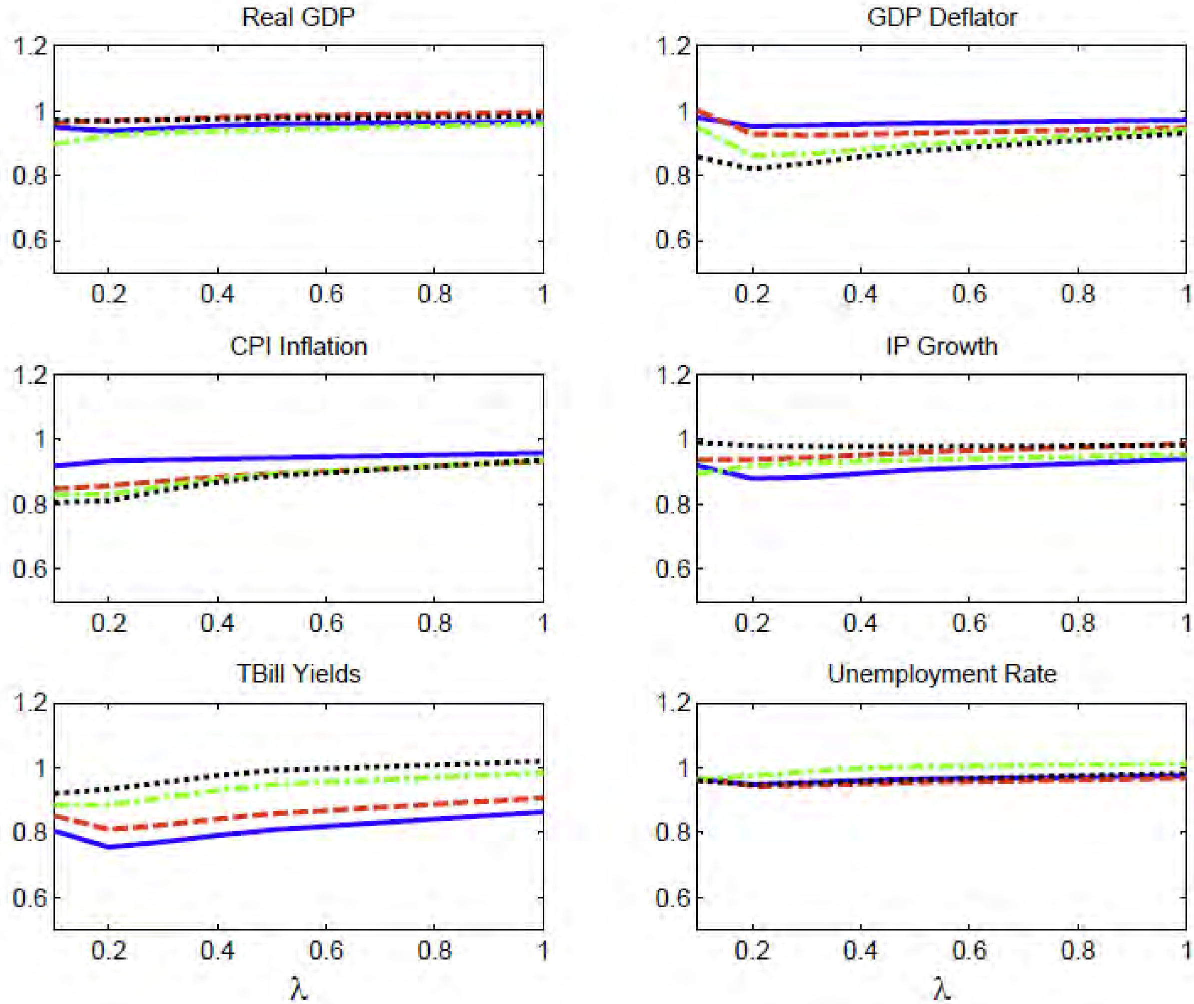
Notes: This table reports the forecast root-mean-square error (RMSE) from estimation of the three-variable VAR of Primiceri (2005) allowing for time-varying parameters and stochastic volatility and from estimation of the ten-variable VAR using the democratic prior D3 with λ=0.1 and λ_0=0.05. In the VAR with time-varying parameters, the three variables are the GDP deflator inflation, the unemployment rate and three-month Treasury Bill yields, and the priors are set following Primiceri. The sample period is 1984Q2-2009Q1. For each horizon and macroeconomic variable, a Diebold-Mariano test of the hypothesis that the two forecasts have equal population RMSE was conducted. One, two, and three asterisks denote cases in which this test rejected the null hypothesis of equal forecast accuracy at the 10, 5, and 1 percent significance levels, respectively.

Table 7: RMSE from the Rolling VAR Estimated by OLS and VAR with democratic Prior

Horizon (Quarters)		h=0	h=1	h=4	h=8
Real GDP Growth	VAR: Dem. Prior	2.09	2.34	2.33	2.34
	Rolling VAR	2.32	2.22	2.29	2.50
GDP Deflator Inflation	VAR: Dem. Prior	1.11	1.07***	1.13**	1.33**
	Rolling VAR	1.22	1.39	1.80	2.07
CPI Inflation	VAR: Dem. Prior	2.05*	2.12**	2.13**	2.19**
	Rolling VAR	2.38	2.51	2.91	3.10
IP Growth	VAR: Dem. Prior	3.98	4.84	5.00	4.89
	Rolling VAR	4.08	4.85	4.96	5.18
TBill Yields	VAR: Dem. Prior	0.62**	0.95***	1.71**	2.29***
	Rolling VAR	0.91	1.36	2.51	3.33
Unemployment Rate	VAR: Dem. Prior	0.30	0.46	0.80	1.00
	Rolling VAR	0.27	0.40	0.75	1.13

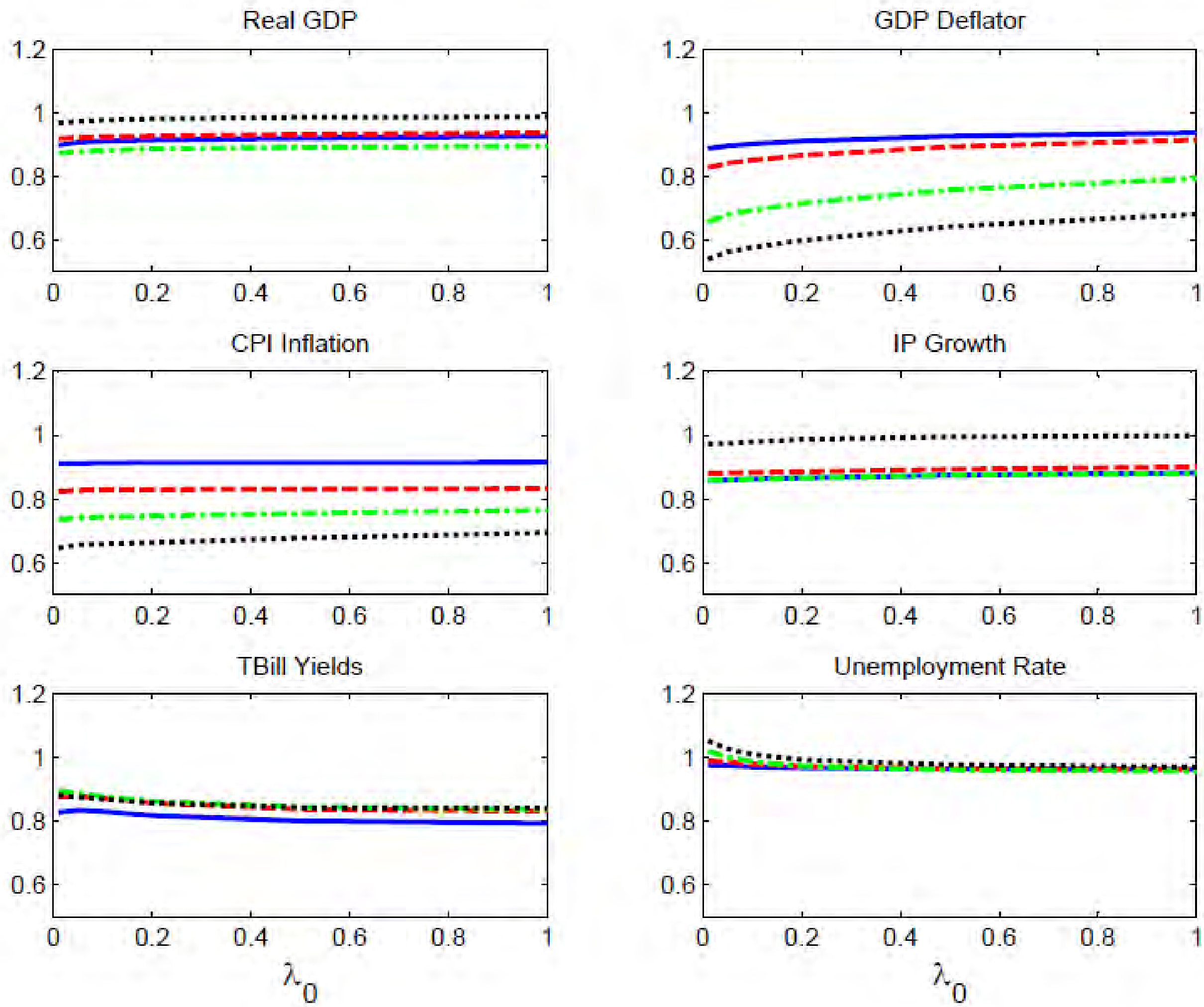
Notes: This table reports the forecast root-mean-square error (RMSE) from OLS estimation of a VAR(1) always using the most recent 40 quarters of data from the real-time dataset for estimation, relative to the RMSE from OLS estimation of a VAR(1) using the full real-time sample. The VAR includes ten variables, as described in the text. The sample period is 1984Q2-2009Q1. For each horizon and macroeconomic variable, a Diebold-Mariano test of the hypothesis that the two forecasts have equal population RMSE was conducted. One, two, and three asterisks denote cases in which this test rejected the null hypothesis of equal forecast accuracy at the 10, 5, and 1 percent significance levels, respectively.

Figure 1: Relative Root-Mean-Square Prediction Errors in VAR using Minnesota Prior (Relative to OLS-based Forecasts)



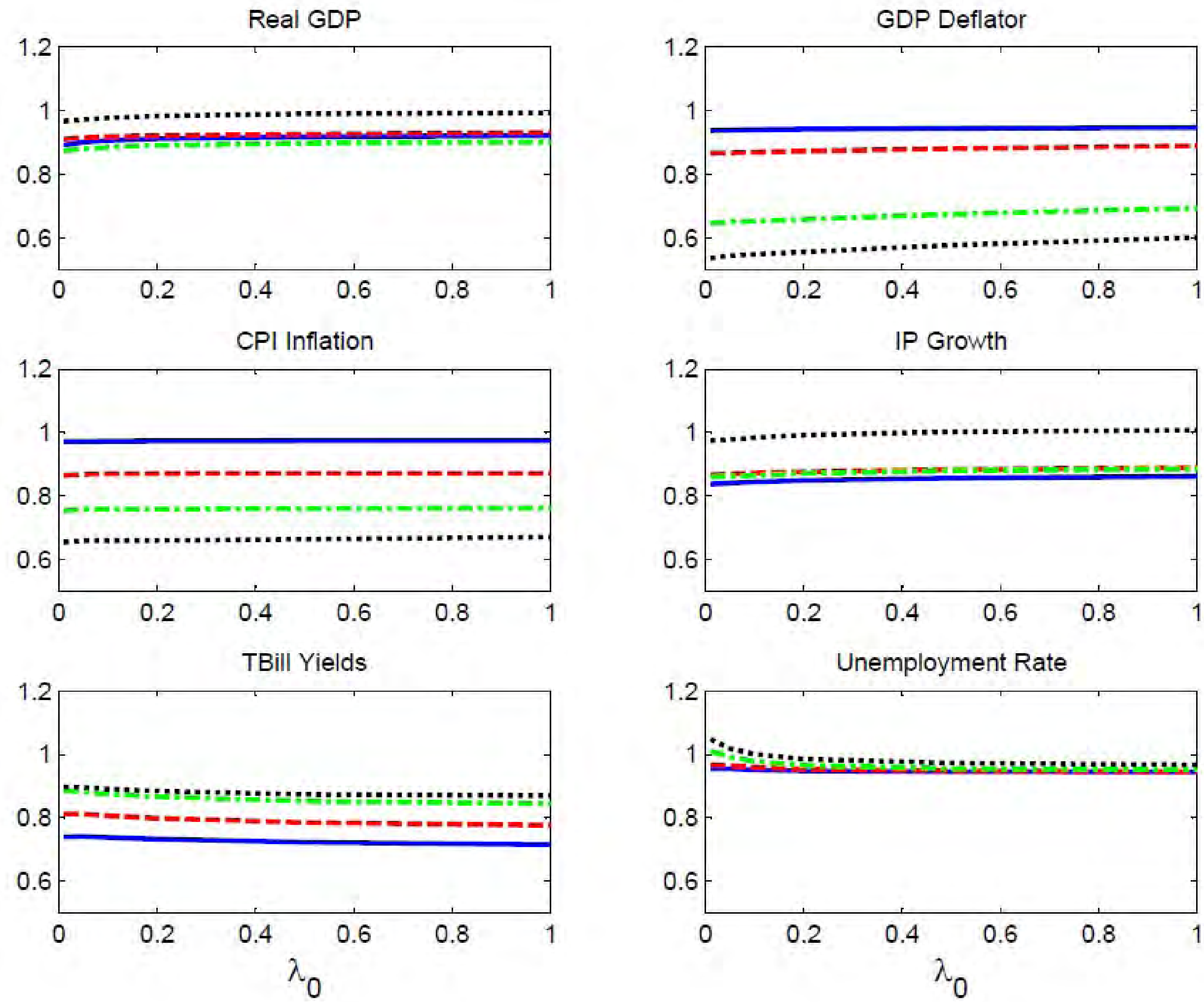
Solid blue line: current-quarter forecast. Dashed red line: one-quarter-ahead forecast. Green dots and dashes: four-quarter-ahead forecast. Black dots: eight-quarter-ahead forecast. The figures plot the ratio of the root-mean-square prediction error of the VAR-based forecasts using the Minnesota prior to that using the OLS estimates against the shrinkage parameter λ. The sample period is 1984Q2-2009Q1 and the VAR contains ten variables, as described in the text.

Figure 2: Relative Root-Mean-Square Prediction Errors in VAR Using democratic Prior D1 (Relative to OLS-based Forecasts)



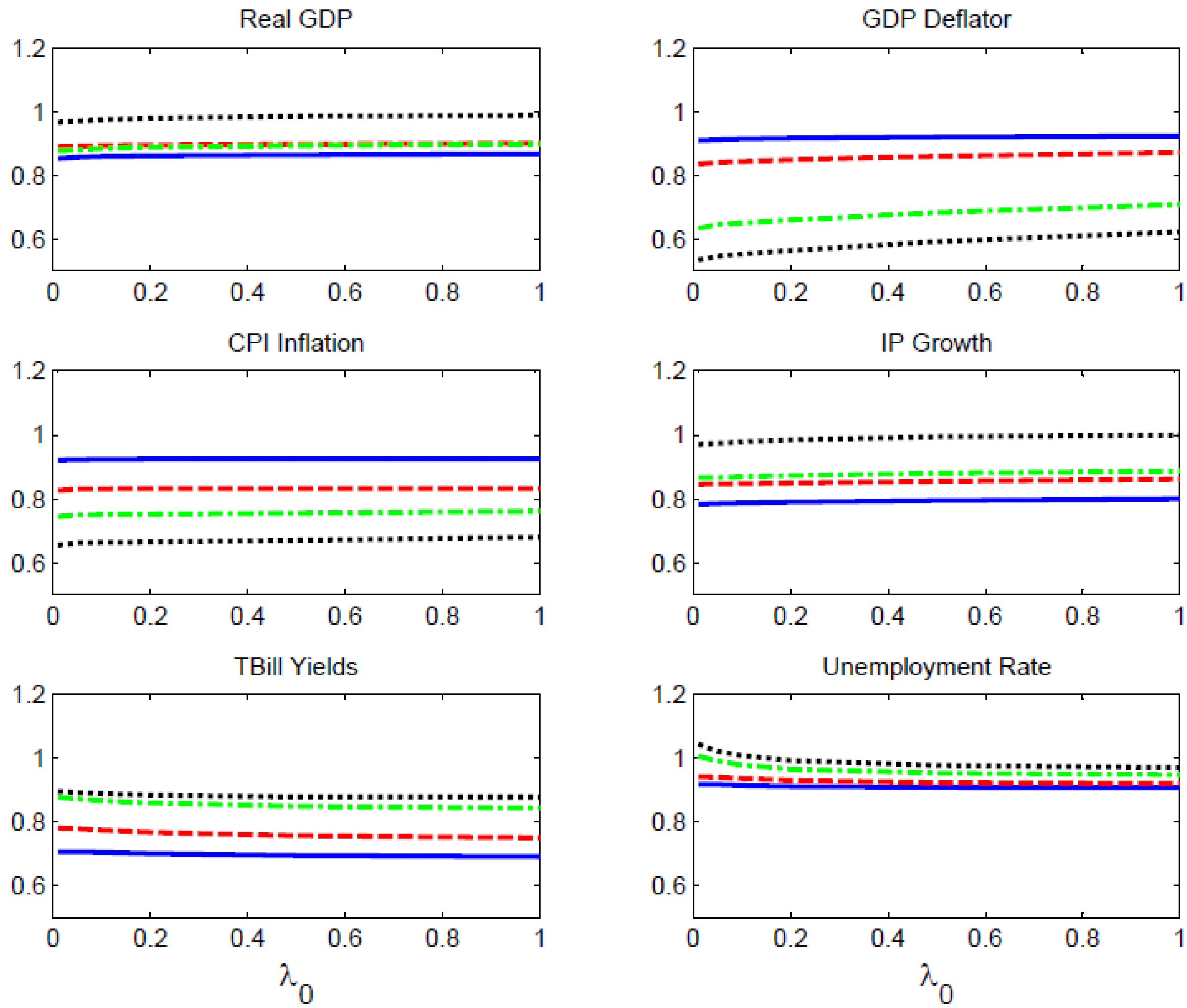
Solid blue line: current-quarter forecast. Dashed red line: one-quarter-ahead forecast. Green dots and dashes: four-quarter-ahead forecast. Black dots: eight-quarter-ahead forecast. The figures plot the ratio of the root-mean-square prediction error of the VAR-based forecasts using the democratic prior D1 to that using the OLS estimates against the shrinkage parameter λ_0. The sample period is 1984Q2-2009Q1, the value of λ is 0.1, and the VAR contains ten variables, as described in the text.

Figure 3: Relative Root-Mean-Square Prediction Errors in VAR Using democratic Prior D2 (Relative to OLS-based Forecasts)



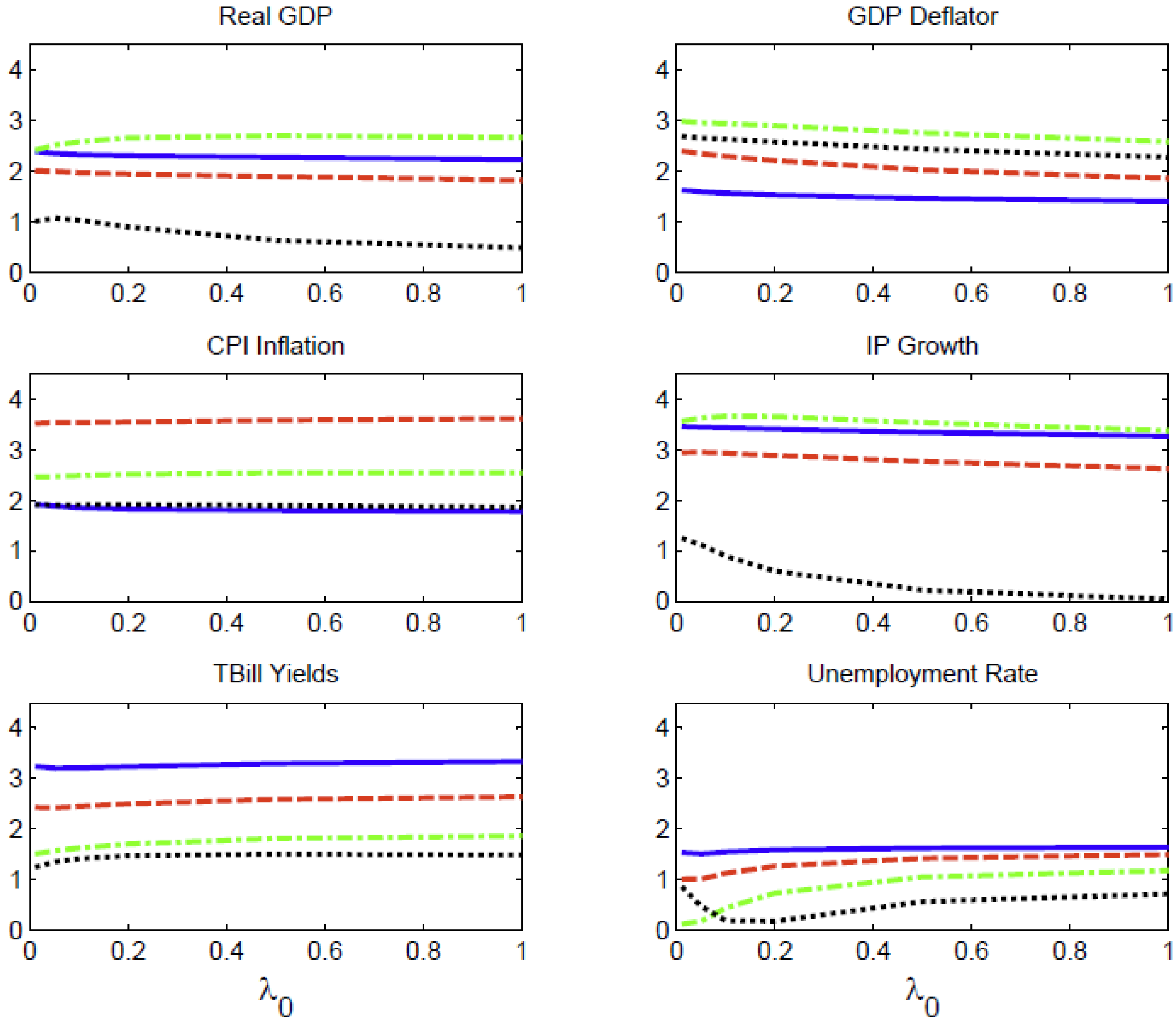
Solid blue line: current-quarter forecast. Dashed red line: one-quarter-ahead forecast. Green dots and dashes: four-quarter-ahead forecast. Black dots: eight-quarter-ahead forecast. The figures plot the ratio of the root-mean-square prediction error of the VAR-based forecasts using the democratic prior D2 to that using the OLS estimates against the shrinkage parameter λ_0. The sample period is 1984Q2-2009Q1, the value of λ is 0.1, and the VAR contains ten variables, as described in the text.

Figure 4: Relative Root-Mean-Square Prediction Errors in VAR Using democratic Prior D3 (Relative to OLS-based Forecasts)



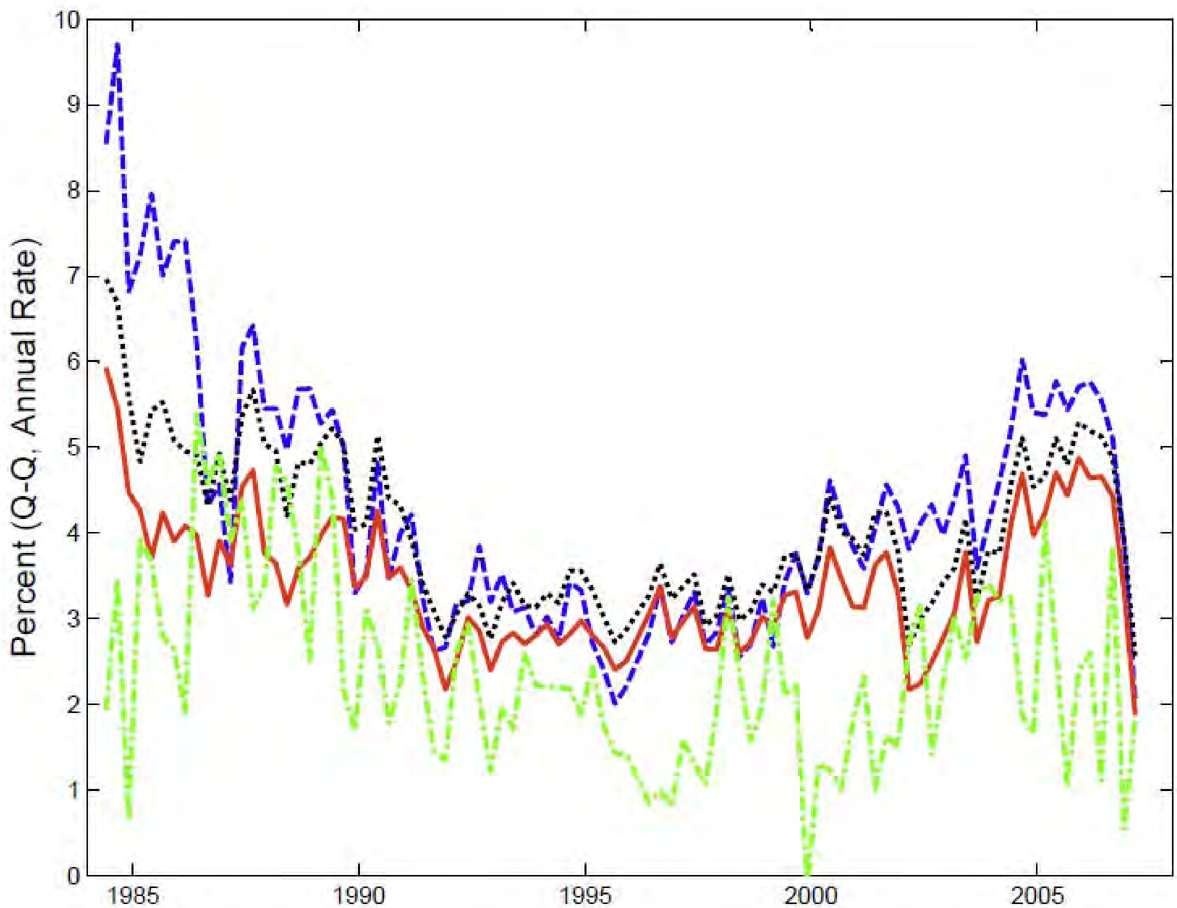
Solid blue line: current-quarter forecast. Dashed red line: one-quarter-ahead forecast. Green dots and dashes: four-quarter-ahead forecast. Black dots: eight-quarter-ahead forecast. The figures plot the ratio of the root-mean-square prediction error of the VAR-based forecasts using the democratic prior D3 to that using the OLS estimates against the shrinkage parameter λ_0. The sample period is 1984Q2-2009Q1, the value of λ is 0.1, and the VAR contains ten variables, as described in the text.

Figure 5: Diebold-Mariano t-statistics testing Equality of Mean Square Prediction Errors Using democratic Prior D3 versus OLS-based Forecasts



Solid blue line: current-quarter forecast. Dashed red line: one-quarter-ahead forecast. Green dots and dashes: four-quarter-ahead forecast. Black dots: eight-quarter-ahead forecast. The figures show the t-statistic proposed by Diebold and Mariano (1995) testing the hypothesis that the mean square prediction errors are equal using the democratic prior and OLS-based forecasts. The t-statistics are plotted against the shrinkage parameter for the democratic prior, λ_0. The sample period is 1984Q2-2009Q1, the value of λ is 0.1, and the VAR contains ten variables, as described in the text.

Figure 6: Time Series of Forecasts of GDP Deflator Inflation Eight Quarters Hence



Dashed blue line: OLS-based forecast. Dotted black line: forecast using the Minnesota prior. Solid red line: forecast using democratic prior D3. Green dashes and dots: actual realized inflation (eight quarters later). The values of the hyperparameters λ and λ_0 are set to 0.1 and 0.05, respectively. The sample period starts in 1984Q2 and the VAR contains ten variables, as described in the text. The forecasts are shown as of the forecast date: The last forecast shown was made in 2007Q1, and is the prediction for 2009Q1.