Research Division
Federal Reserve Bank of St. Louis
Working Paper Series

Evaluating Conditional Forecasts from Vector Autoregressions

Todd E. Clark
And
Michael W. McCracken

Working Paper 2014-025B
http://research.stlouisfed.org/wp/2014/2014-025.pdf

September 2014
Revised April 2015

FEDERAL RESERVE BANK OF ST. LOUIS
Research Division
P.O. Box 442
St. Louis, MO 63166

The views expressed are those of the individual authors and do not necessarily reflect official positions of the Federal Reserve Bank of St. Louis, the Federal Reserve System, or the Board of Governors.

Federal Reserve Bank of St. Louis Working Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to Federal Reserve Bank of St. Louis Working Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

Evaluating Conditional Forecasts from Vector Autoregressions *

Todd E. Clark
Federal Reserve Bank of Cleveland

Michael W. McCracken
Federal Reserve Bank of St. Louis

April 2015

Abstract

Many forecasts are conditional in nature. For example, a number of central banks routinely report forecasts conditional on particular paths of policy instruments. Even though conditional forecasting is common, there has been little work on methods for evaluating conditional forecasts. This paper provides analytical, Monte Carlo, and empirical evidence on tests of predictive ability for conditional forecasts from estimated models. In the empirical analysis, we consider forecasts of GDP and investment growth as well as inflation from a VAR, based on conditions on the short-term interest rate. Throughout the analysis, we focus on tests of bias, efficiency, and equal accuracy applied to conditional forecasts from VAR models.

JEL Nos.: C53, C52, C12, C32
Keywords: Prediction, forecasting, out-of-sample

* *Clark*: Economic Research Dept.; Federal Reserve Bank of Cleveland; P.O. Box 6387; Cleveland, OH 44101; `todd.clark@clev.frb.org`. *McCracken* (corresponding author): Research Division; Federal Reserve Bank of St. Louis; P.O. Box 442; St. Louis, MO 63166; `michael.w.mccracken@stls.frb.org`. The authors gratefully acknowledge helpful comments from Graham Elliott, Ana Galvao, Silvia Goncalves, James Nason, Frank Schoerfheide, Dan Waggoner, Mark Watson, and participants at the 2014 NBER Summer Institute, 2014 Barcelona GSE Summer Forum on Time Series Analysis, the IAAE 2014 Annual Conference, the $(EC)^2$ 2014 Annual Conference, and the 2014 EABCN conference on Judgement and Combination in Forecasting and Policy Models. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, Federal Reserve Bank of St. Louis, Federal Reserve System, or any of its staff.

1 Introduction

Since the seminal work of Litterman (1986), vector autoregressions (VARs) have been widely known to be useful for out-of-sample forecasting. In many applications, the forecasts of interest are unconditional. Clark and McCracken (2013) review methods for evaluating such forecasts.

VARs are also commonly used to construct conditional forecasts. In such applications, the models are used to predict variables such as GDP growth and inflation conditional on, e.g., an assumed path of monetary policy variables or an assumed path of oil prices. Examples of VAR forecasts conditional on policy paths include Sims (1982), Doan, Litterman, and Sims (1984), and Meyer and Zaman (2013). Giannone, et al. (2014) use VARs to construct forecasts of inflation conditional on paths for oil and other price indicators. Baumeister and Kilian (2013) consider forecasts of oil prices conditioned on a range of scenarios. Schorfheide and Song (2013) and Aastveit, et al. (2014) use VARs to produce multi-step forecasts of growth, inflation, and other macroeconomic variables conditional on current-quarter forecasts obtained from other, judgmental sources (the Federal Reserve's Greenbook for the former and Survey of Professional Forecasters for the latter).

In light of common need for conditional forecasts, one would like to have a feel for their quality. Accordingly, in this paper, we develop and apply methods for the evaluation of conditional forecasts from VARs, using tests of bias, efficiency, and the MSE accuracy of conditional versus unconditional forecasts. More specifically, we provide analytical, Monte Carlo, and empirical evidence on tests of predictive ability for conditional forecasts from estimated models. In the empirical analysis, we consider forecasts of growth, fixed investment, and inflation from a VAR, based on conditions on the short-term interest rate.

Throughout, our intention is to provide usable metrics for evaluating conditional forecasts, in a general sense and in comparison to the accuracy of unconditional forecasts. To do so, we focus on particular forms of conditional forecasts for which interpretation of various null and alternative hypotheses is most straight-forward. In particular, in our analysis, we consider forecasts conditioned on actual future information on some variables in the VAR model. In practice, conditional forecasts are sometimes constructed with future information (e.g., based on forward guidance from the central bank about policy rates), but not always. In general, though, the efficacy of conditional forecasts rests on having a properly specified model. Our testing based on future information provides a way of assessing proper

specification of the VAR. As we detail below, Herbst and Schorfheide (2012) use a similar idea in a Bayesian evaluation framework.

To better understand our approach to inference, consider a very simple example in which we forecast inflation (y_t) conditioned on a path for the federal funds rate (x_t) over the next two periods $t+1$ and $t+2$. Suppose that the data-generating process for inflation and the funds rate is a zero-mean stationary VAR(1) taking the form

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} e_t \\ v_t \end{pmatrix},$$

with N(0,1) errors with contemporaneous correlation ρ. Following the approach taken in Doan, Litterman, and Sims (1984), conditional on this path for the funds rate, the minimum mean square error (MSE) one- and two-step ahead forecasts of y_t are as follows:

$$\begin{aligned} \hat{y}^c_{t,1} &= \hat{y}^u_{t,1} + \hat{\rho}(\hat{x}^c_{t,1} - \hat{x}^u_{t,1}) \\ \hat{y}^c_{t,2} &= \hat{y}^u_{t,2} + (\hat{b} + \hat{\rho}(\hat{a} - \hat{c}))(\hat{x}^c_{t,1} - \hat{x}^u_{t,1}) + \hat{\rho}(\hat{x}^c_{t,2} - \hat{x}^u_{t,2}), \end{aligned}$$

where the superscripts c and u denote conditional and unconditional forecasts, respectively. In both cases the conditional forecasts of y are comprised of the standard, unconditional MSE-optimal forecast $\hat{y}^u_{t,j}$, $j = 1, 2$, plus additional terms that capture the impact of conditioning on future values of the federal funds rate, $\hat{x}^c_{t,1}$ and $\hat{x}^c_{t,2}$.

After rearranging terms, the conditional forecast errors $\hat{\varepsilon}^c_{t,2} = y_{t+2} - \hat{y}^c_{t,2}$ and $\hat{\varepsilon}^c_{t,1} = y_{t+1} - \hat{y}^c_{t,1}$ take the form

$$\begin{aligned} \hat{\varepsilon}^c_{t,1} &= \hat{\varepsilon}^u_{t,1} - \hat{\rho}(\hat{v}^u_{t,1} - \hat{v}^c_{t,1}) \\ \hat{\varepsilon}^c_{t,2} &= \hat{\varepsilon}^u_{t,2} - (\hat{b} + \hat{\rho}(\hat{a} - \hat{c}))(\hat{v}^u_{t,1} - \hat{v}^c_{t,1}) - \hat{\rho}(\hat{v}^u_{t,2} - \hat{v}^c_{t,2}). \end{aligned}$$

We immediately see that any "good" properties that the conditional forecast errors $\hat{\varepsilon}^c_{t,1}$ and $\hat{\varepsilon}^c_{t,2}$ have, such as unbiasedness or efficiency, are jointly determined by: (i) the quality of the unconditional forecast errors $\hat{\varepsilon}^u_{t,1}$, $\hat{\varepsilon}^u_{t,2}$, $\hat{v}^u_{t,1}$, and $\hat{v}^u_{t,2}$, as well as (ii) the behavior of the conditioning as measured via $\hat{v}^c_{t,1}$ and $\hat{v}^c_{t,2}$. As such, any method of inference designed to evaluate the quality of the conditional forecasts must somehow distinguish between properties determined by the quality of the model and properties determined by the quality of the conditioning.

Because these properties are difficult to separate in such a way as to yield useful tests of, for example, forecast efficiency, we consider tests based on a more tractable approach

in which we condition on the ex-post realized values of the conditioning variables so that $\hat{x}^c_{t,1} = x_{t+1}$ and $\hat{x}^c_{t,2} = x_{t+2}$. Since both $\hat{v}^c_{t,1}$ and $\hat{v}^c_{t,2}$ are then numerically zero we find that any properties associated with the conditional forecast errors $\hat{\varepsilon}^c_{t,1}$ and $\hat{\varepsilon}^c_{t,2}$ are determined only by the quality of the VAR as measured though the unconditional forecast errors, the regression parameters, and the residual variance parameters. Because we take this approach to conditioning, our inferential procedures are designed to evaluate the ability of the VAR to construct good conditional forecasts rather than evaluating a specific set of conditional forecasts per se.

Because we take this approach to inference, one might wonder how our tests differ from those already in the literature that evaluate the quality of unconditional forecasts. The primary difference can be seen in the simple example above. Properties of a good conditional forecast depend not only on the properties of a good unconditional forecast but also correct specification of the residual variance matrix (in this case via ρ). As we will see later, certain tests of efficiency and equal accuracy, when applied to conditional forecasts, have power in directions associated with misspecification in the residual variance that comparable tests applied to unconditional forecasts do not. That said, we make no claim that the tests delineated in this paper encompass existing ones targeting unconditional forecasts. Misspecification in the lagged (rather than contemporary) dynamics of the VAR often yield higher power in tests that focus on unconditional, rather than conditional, forecasts. As such we view our results as a complement to the existing literature - a complement that emphasizes that good unconditional forecasting need not imply good conditional forecasting.

Our work builds upon two previous studies that have developed methods for the evaluation of some form of conditional forecasts. Motivated by an interest in evaluating the efficiency of Greenbook forecasts, Faust and Wright (2008) develop a regression-based test of predictive ability that accounts for the conditioning of the Greenbook forecasts on a pre-specified path of the federal funds rate over the forecast horizon. For the purpose of evaluating forecasts from DSGE models, Herbst and Schorfheide (2012) develop Bayesian methods to check the accuracy of point and density forecasts. More specifically, Herbst and Schorfheide consider the Bayesian tool of posterior predictive checks and forecasts of each variable conditioned on the actual future path of another, selected variable. Our paper differs from these in that we focus on conditional forecasts from VARs and emphasize frequentist inference.

The remainder of the paper proceeds as follows. Section 2 describes the approach to conditional forecasting that we consider. Section 3 provides our theoretical results (proofs are provided in the Appendix). Section 4 presents Monte Carlo evidence on the finite-sample accuracy of our proposed methods for evaluating conditional forecasts from VARs. Section 5 presents a practical application, to macroeconomic forecasts conditioned on interest rate paths. Section 6 concludes.

2 Conditional Forecasting Approaches

In generating and evaluating conditional forecasts, we focus on the standard approach in VAR forecasting, which we refer to as the minimum-MSE approach. This standard is based on the textbook problem of conditional projection, as could be handled with a state space formulation of the VAR and the Kalman filter and smoother (see, e.g., Clarida and Coyle (1984) or Giannone, et al. (2014)). The conditions on the variables of interest are contained in the measurement vector and equation; the data vector of the VAR is the state vector of the transition equation. The projection problem is one of predicting the state vector given the measurements (conditions). Doan, Litterman, and Sims (1984) developed an alternative approach to solving this formulation of the conditional forecasting problem, which consists of solving a least squares problem to pick the shocks needed to satisfy the conditions. In the context of conditioning on a policy path, this approach to conditional forecasting can be seen as consisting of the following: determining the set of shocks to the VAR that, by a least squares metric, best meet the conditions on the policy rate. In practice, this approach may mean that, in a given episode, an unchanged path of the policy rate could, for example, be due to shocks to output. Under this approach, the conditional forecasts are **not** dependent on the identification of structural shocks in the VAR. Note that under this approach, the forecast for each period can be affected by the imposition of conditions at all periods. For example, if we impose conditions for two periods, the forecast for period $t+2$ will generally be affected by the conditions on both period $t+1$ and $t+2$.

In DSGE forecasting, the more standard approach for achieving conditions on the policy path rests on feeding in structural shocks (either anticipated or unanticipated) to monetary policy needed to hit the policy path. For example, Del Negro and Schorfheide (2013, p.62) assess "...using unanticipated and anticipated monetary policy shocks to generate forecasts conditional on a desired interest rate path." In some cases, a combination of shocks to

policy and other structural shocks is used to hit the desired policy path.

In a VAR setting, of course, one does not have the option of using anticipated policy shocks to achieve a policy path. One might instead use a sequence of unanticipated policy shocks to achieve the desired path of the policy variable, under a chosen scheme for identifying policy shocks, based on the approaches of studies such as Bernanke and Blinder (1992), Christiano, Eichenbaum, and Evans (1996), and Giannone, Lenza, and Primiceri (2012). In earlier versions of this paper, we included such a policy shock-based approach, and we obtained results very similar to those we report for the minimum-MSE approach more typically used for conditional forecasting with VARs.

3 Analytical results

We present our theoretical results in an environment in which OLS-estimated VARs are used to construct τ-step ahead conditional forecasts sequentially across forecast origins $t = R, ..., T - \tau = R + P - \tau$ for a fixed and finite horizon τ. Specifically, suppose that the model takes the form

$$Y_t = C + A(L)Y_{t-1} + \varepsilon_t,$$

where $Y = (y_1, y_2, ..., y_n)'$, $\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_n)'$, and $A(L) = \sum_{j=1}^{l} A_j L^j$ for $n \times 1$ and $n \times n$ parameter matrices C and A_j, $j = 1, ..., l$, respectively. This is equivalent to

$$Y_t = \Lambda x_{t-1} + \varepsilon_t = (x'_{t-1} \otimes I_n)\beta + \varepsilon_t$$

if we define $x_t = (1, Y'_t, ..., Y'_{t-l+1})'$, $\beta = vec(\Lambda)$, and $\Lambda = (C, A_1, ..., A_l)$. If the model is estimated using the recursive scheme we obtain the estimated regression parameters

$$\hat{\beta}_t = vec(\hat{\Lambda}_t) = vec((t^{-1} \sum_{s=1}^{t-1} Y_{s+1} x'_s)(t^{-1} \sum_{s=1}^{t-1} x_s x'_s)^{-1})$$

and corresponding residual variance matrix

$$\hat{\Sigma}_t = t^{-1} \sum_{s=1}^{t-1} \hat{\varepsilon}_{s+1} \hat{\varepsilon}'_{s+1}.$$

If the model is estimated using a rolling window of R observations we obtain comparable estimators but defined over the observations $s = t - R + 1, ..., t$ rather than $s = 1, ..., t$.

Both the regression and variance estimates $\hat{\beta}_t$ and $\hat{\Sigma}_t$ are used to construct iterated multi-step unconditional and conditional forecasts of Y_{t+j}, denoted $\hat{Y}^u_{t,j}$ and $\hat{Y}^c_{t,j}$, which in turn imply forecast errors $\hat{\varepsilon}^u_{t,j}$ and $\hat{\varepsilon}^c_{t,j}$. If we are interested in forecasting the ith element of

$Y_{t+\tau}$, define $\hat{y}_{i,t,\tau} = \iota_i' \hat{Y}_{t,\tau}$ for the vector ι_i with a 1 in the ith position and zeroes elsewhere. Define $\hat{\varepsilon}^c_{i,t,j} = \iota_i' \hat{\varepsilon}^c_{t,j} = y_{i,t+j} - \hat{y}^c_{i,t,j}$ and $\hat{\varepsilon}^u_{i,t,j} = \iota_i' \hat{\varepsilon}^u_{t,j} = y_{i,t+j} - \hat{y}^u_{i,t,j}$, accordingly. To simplify notation, and without loss of generality, we assume that the first element of Y is the primary element of interest and hence most of our analysis emphasizes the properties of the conditional forecast $\hat{y}^c_{1,t,j}$ and associated forecast error $\hat{\varepsilon}^c_{1,t,j}$.

As discussed in Section 2, the conditional forecasts are assumed to be obtained by the standard minimum-MSE approach. The resulting forecasts take the form

$$\hat{y}^c_{1,t,\tau} = \hat{y}^u_{1,t,\tau} + \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{m} \hat{\gamma}_{i,t,j} (\hat{y}^u_{i,t,j} - \hat{y}^c_{i,t,j})$$

for a collection of constants $\hat{\gamma}_{i,j}$ that are non-stochastic functions of both $\hat{\beta}_t$ and $\hat{\Sigma}_t$.[1] Note that this structure aligns with the simple example from the introduction. In addition, we allow the maximum conditioning horizon m to be greater than or less than the forecast horizon τ, and we allow for direct conditioning on some elements of the future path of y_1 itself.

As mentioned earlier, we focus on evaluating the ability of the model to construct good conditional forecasts rather than evaluating conditional forecasts per se. We do so by examining the properties of the conditional forecast error when we condition on future realized values of those variables in the hypothetical scenario of interest. Specifically, if we set $\hat{y}^c_{i,t,j} = y_{i,t+j}$ we obtain

$$\hat{\varepsilon}^c_{1,t,\tau} = \hat{\varepsilon}^u_{1,t,\tau} - \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{m} \hat{\gamma}_{i,t,j} \hat{\varepsilon}^u_{i,t,j}.$$

By taking this approach, the conditional forecast error has a representation as a linear function of unconditional forecast errors across all variables in the scenario and across all conditioning horizons.

Before proceeding to the tests we need to introduce some notation. Define $\hat{\phi}_t = (\hat{\beta}_t', vech(\hat{\Sigma}_t)')'$, $\phi = (\beta', vech(\Sigma)')'$, $B_\beta = ((Ex_s x_s')^{-1} \otimes I_n)$, $B_\Sigma = I_{n(n+1)/2}$, $B = diag(B_\beta, B_\Sigma)$, and $h_{s+1} = (h'_{\beta,s+1}, h'_{\Sigma,s+1})' = (vec(\varepsilon_{s+1} x_s')', vech(\varepsilon_{s+1}\varepsilon'_{s+1} - \Sigma)')'$. For any parametric function $z_t(\cdot)$, let $\hat{z}_t = z_t(\hat{\phi}_t)$ and $z_t = z_t(\phi)$. Finally, the asymptotics we use require both the initial in-sample size R and number of forecasts P to diverge as the overall sample increases. By taking this approach the asymptotic distribution is influenced by the ratio

[1] As a practical matter, many of the $\gamma_{i,j}$ will be zero depending on how many variables are conditioned on and how long the maximal conditioning horizon is. In the assumptions, we impose the restriction that $\gamma_{1,\tau}$ is zero (and hence the value of $\hat{y}^c_{1,t,\tau}$ is not imposed directly) via rank conditions on certain variance matrices.

$\lim_{P,R\to\infty} P/R = \pi$. In particular, the asymptotic variances developed below all depend on the weights λ_{fh} and λ_{hh} as described in the following table taken from West and McCracken (1998).

	$\lambda_{fh} =$	$\lambda_{hh} =$
Recursive	$1 - \pi^{-1}\ln(1+\pi)$	$2(1 - \pi^{-1}\ln(1+\pi))$
Rolling, $\pi \leq 1$	$\pi/2$	$\pi - \pi^2/3$
Rolling, $1 < \pi < \infty$	$1 - (2\pi)^{-1}$	$1 - (3\pi)^{-1}$

3.1 Regression based tests of bias and efficiency

In this section we develop tests of zero bias and efficiency in the context of conditional forecasts from VARs when conditioning on future values of the variables in a hypothetical scenario. To do so first note that each can be couched in the context of a test of the null hypothesis that the coefficient α is zero in the regression

$$\hat{\varepsilon}^c_{1,t,\tau} = \hat{g}_t'\alpha + error,\ t = R, ..., T - \tau$$

for appropriate definitions of $\hat{g}_t$. Examples include a test of zero bias if we let $\hat{g}_t = 1$ and a test of efficiency if we let $\hat{g}_t = (1, \hat{y}^u_{1,t,\tau})$ or perhaps $\hat{g}_t = (1, \hat{y}^u_{1,t,\tau}, \hat{y}^u_{i,t,j})$ for those variables y_i in the scenario conditioned at horizon j.[2]

In each case, $P^{1/2}\hat{\alpha}$ is asymptotically normal with zero mean and a variance that accounts for estimation error in the estimated conditional forecast errors $\hat{\varepsilon}^c_{1,t,\tau}$ and generated regressors $\hat{g}_t$. These results follow directly from Lemmas 4.1 - 4.3 and Theorems 4.1 and 4.2 of West and McCracken (1998) if, under the null that $\alpha = 0$, we maintain a correctly specified VAR with errors ε_t that form a martingale difference sequence. Nevertheless it is worth noting that the *interpretation* of the results is slightly different than that intended in West and McCracken (1998). This arises because the conditional forecast errors depend on $\hat{\phi}_t = (\hat{\beta}_t', vech(\hat{\Sigma}_t)')'$ rather than just $\hat{\beta}_t$. Nevertheless, $\hat{\phi}_t$ satisfies assumption A2 of West and McCracken (1998) and the remainder of their assumptions are satisfied except for A1 (c) and (d).[3] These assumptions, however, do not affect the derivation of asymptotic normality. They are only used in Theorem 5.1 of West and McCracken (1998), wherein special cases are delineated under which estimation error is asymptotically irrelevant. For clarity we restate the assumptions in the context of the current paper and then proceed to

[2] Note that this last regression is not the standard Mincer-Zarnowitz form of the efficiency regression. We consider this separately in the next section.

[3] See section 3.4.3 of Lutkepohl (1991) for details.

the Theorem.[4]

Assumption 1: (a) In some neighborhood N around ϕ, and with probability 1, $\varepsilon^c_{1,t,\tau}(\phi)$ and $g_t(\phi)$ are measurable and twice continuously differentiable; (b) $E(\varepsilon_t|x_{t-j}, \varepsilon_{t-j}$ all $j \geq 1) = 0$; (c) Eg_tg_t' is full rank.

Assumption 2: The estimate $\hat{\phi}_t$ satisfies $\hat{\phi}_t - \phi = B(t)H(t)$, with (a) $B(t) \rightarrow_{a.s.} B = diag[((Ex_sx_s')^{-1} \otimes I_n), I_{n(n+1)/2}]$ and B of full rank; (b) $H(t)$ equals $t^{-1}\sum_{s=1}^{t-1} h_{s+1}$ and $t^{-1}\sum_{s=t-R+1}^{t-1} h_{s+1}$ for the recursive and rolling schemes, respectively; (c) $Eh_{s+1} = 0$; (d) In the neighborhood N of Assumption 1, h_{s+1} is measurable and continuously differentiable.

Assumption 3: In the neighborhood N of Assumption 1, there is a constant $D < \infty$ such that for all t, $\sup_{\phi \in N} |\partial^2 \varepsilon^c_{1,t,\tau}(\phi)/\partial\phi\partial\phi'| < m_t$ for a measurable m_t for which $Em_t^4 < D$. The same holds when $\varepsilon^c_{1,t,\tau}(\phi)$ is replaced by $g_t(\phi)$.

Assumption 4: Let $w_t = (x_t', vec(\partial g_t/\partial\phi)', \varepsilon_t', g_t', h_t')'$. (a) For some $d > 1$, $\sup_t E||w_t||^{8d} < \infty$, where $||\cdot||$ denotes the Euclidean norm; (b) w_t is strong mixing with coefficients of size $-3d/(d-1)$; (c) w_t is fourth-order stationary; (d) $S_{ff} = \lim_{P,R\rightarrow\infty} Var(P^{-1/2}\sum_{t=R}^{T-\tau} \varepsilon^c_{1,t,\tau} g_t)$ is positive definite.

Assumption 5: $R, P \rightarrow \infty$ as $T \rightarrow \infty$ with $\lim_{T\rightarrow\infty} \frac{P}{R} = \pi$. (a) $0 \leq \pi \leq \infty$ for the recursive scheme; (b) $0 \leq \pi < \infty$ for the rolling scheme.

Theorem 1 *Maintain assumptions A1-A5.* $P^{1/2}\hat{\alpha} \rightarrow^d N(0, V)$ *with* $V = (Eg_tg_t')^{-1}\Omega(Eg_tg_t')^{-1}$ *and*

$$\Omega = S_{ff} + \lambda_{fh}(FBS_{fh}' + S_{fh}B'F') + \lambda_{hh}FBS_{hh}B'F',$$

where $S_{ff} = \lim_{P,R\rightarrow\infty} Var(P^{-1/2}\sum_{t=R}^{T-\tau} \varepsilon^c_{1,t,\tau} g_t)$, $S_{hh} = \lim_{T\rightarrow\infty} Var(T^{-1/2}\sum_{s=1}^{T-1} h_{s+1})$, $S_{fh} = \lim_{P,R\rightarrow\infty} Cov(P^{-1/2}\sum_{t=R}^{T-\tau} \varepsilon^c_{1,t,\tau} g_t, P^{-1/2}\sum_{t=R}^{T-\tau} h_{t+1}')$, *and* $F = E(\partial\varepsilon^c_{1,t,\tau}(\phi)g_t(\phi)/\partial\phi)$.

As in West and McCracken (1998), the asymptotic variance, especially through Ω, is comprised of three components: S_{ff} captures the variation that would exist even if ϕ were known, $FBS_{hh}B'F'$ captures the variation due purely to estimation error in $\hat{\phi}_t$, and $FBS_{fh}' + S_{fh}B'F'$ is the covariance between the two sources of variability.

The asymptotic variance is complicated but can be simplified for each of the tests discussed above. First, note that since $E\varepsilon^c_{1,t,\tau}(\partial g_t(\phi)/\partial\phi) = 0$, F trivially reduces to

[4] While notationally very similar, these assumptions are stronger than those in West and McCracken (1998). For example, in Assumption 1 measurability and differentiability is imposed on $\varepsilon^c_{1,t,\tau}(\phi)$ rather than $\varepsilon^u_{1,t,\tau}(\phi)$. But since $\varepsilon^c_{1,t,\tau}(\phi)$ is a function of unconditional forecast errors from other equations and horizons, this assumption effectively imposes measurability and differentiability on $\varepsilon^u_{i,t,j}(\phi)$ for all equations $i = 1, ..., n$ and horizons $j = 1, ..., m$.

$Eg_t(\partial\varepsilon^c_{1,t,\tau}(\phi)/\partial\phi)$. In addition, it is straightforward to show that $Eg_t(\partial\varepsilon^c_{1,t,\tau}(\phi)/\partial vech(\Sigma)) = 0$ and hence $F = (Eg_t(\partial\varepsilon^c_{1,t,\tau}(\phi)/\partial\beta), 0)$. That, along with the fact that B is block diagonal, implies that in large samples there is no estimation error contributed by $\hat{\Sigma}_t$ in either $\hat{g}_t$ or $\hat{\varepsilon}^c_{1,t,\tau}$. Notationally, we can therefore simplify the formula for Ω to

$$\Omega = S_{ff} + \lambda_{fh}(F_\beta B_\beta S'_{fh,1} + S_{fh,1}B'_\beta F'_\beta) + \lambda_{hh}F_\beta B_\beta S_{hh,11}B'_\beta F'_\beta,$$

where $F_\beta = Eg_t(\partial\varepsilon^c_{1,t,\tau}(\phi)/\partial\beta)$ and $S_{fh,1}$ and $S_{hh,11}$ denote those elements of S_{fh} and S_{hh} associated only with the OLS moment conditions $h_{\beta,s+1}$.

With a bit more work, F can be explicitly derived using elements of section 3.5.2 of Lutkepohl (1991). Specifically if we define the $(nl+1)\times(nl+1)$ matrix

$$W = \begin{pmatrix} 1 & 0 & \dots & 0 \\ & \Lambda & & \\ 0 & I_{n(l-1)} & & 0 \end{pmatrix},$$

the $n \times (nl+1)$ selection matrix $J_1 = (0_{n\times 1}, I_n, 0_{n\times n(l-1)})$, $\Phi_i = J_1 W^i J'_1$, and $\Theta_{i,t,j} = \iota'_i \sum_{q=0}^{j-1}(x'_t(W')^{j-1-q} \otimes \Phi_q)$, we obtain

$$F_\beta = Eg_t(-\Theta_{1,t,\tau} + \sum\nolimits_{i=1}^{n}\sum\nolimits_{j=1}^{m}\gamma_{i,j}\Theta_{i,t,j}).$$

3.2 Mincer-Zarnowitz efficiency test

In the previous section we delineated a regression-based test of efficiency based on a regression of $\hat{\varepsilon}^c_{1,t,\tau}$ on $(1, \hat{y}^u_{1,t+\tau})$. By taking this approach we were able to use the results in West and McCracken (1998) to prove that the regression coefficients are asymptotically normal under the null hypothesis that $\alpha = 0$. While useful, this result deviates from a more standard version of the Mincer and Zarnowitz (1969) test of efficiency in which, for conditional forecasts, we would regress $\hat{\varepsilon}^c_{1,t,\tau}$ on $(1, \hat{y}^c_{1,t,\tau})$.

The reason we didn't consider the Mincer-Zarnowitz regression is that, by conditioning on future values of the variables in the scenario, it is possible that $E\hat{\varepsilon}^c_{1,t,\tau}\hat{y}^c_{1,t+\tau} \neq 0$ even though $\hat{\varepsilon}^c_{1,t,\tau}$ is orthogonal to all information in the time t information set. However, as it turns out, with minimum-MSE conditional forecasts, that orthogonality restriction holds.[5] Straightforward algebra show that $E\hat{\varepsilon}^c_{1,t,\tau}\hat{y}^c_{1,t+\tau} = 0$ and hence Theorem 1 is indeed applicable with $\hat{g}_t = (1, \hat{y}^c_{1,t,\tau})'$. Even so, the previously discussed simplifications of the asymptotic variance Ω no longer hold. In particular, $\hat{\Sigma}_t$ contributes estimation error even in the limit.

[5] However, if the conditional forecasts are produced by the policy shock approach mentioned in section 2, the orthogonality condition no longer holds.

We can see this directly in the following derivation of the $(2 \times (\dim vec(\Lambda) + \dim vech(\Sigma)))$ matrix $F = (F_1', F_2')' = E(\partial \varepsilon^c_{1,t,\tau}(\phi) g_t(\phi)/\partial\phi)$:

$$
\begin{aligned}
F_1 &= (E(-\Theta_{1,t,\tau} + \sum_{i=1}^n \sum_{j=1}^m \gamma_{i,j}\Theta_{i,t,j}), 0_{n(n+1)/2}) \\
F_2 &= (E(y^u_{1,t,\tau})(-\Theta_{1,t,\tau} + \sum_{i=1}^n \sum_{j=1}^m \gamma_{i,j}\Theta_{i,t,j}) - E(y^c_{1,t,\tau} - y^u_{1,t,\tau})(\sum_{i=1}^n \sum_{j=1}^m \nabla_\beta \gamma_{i,j} u^u_{i,t,j}) \\
&\quad + E(u^c_{1,t,\tau})(\sum_{i=1}^n \sum_{j=1}^m \nabla_\beta \gamma_{i,j} u^u_{i,t,j}), -E(y^c_{1,t,\tau})(\sum_{i=1}^n \sum_{j=1}^m \nabla_\Sigma \gamma_{i,j} u^u_{i,t,j}) \\
&\quad + E(u^c_{1,t,\tau})(\sum_{i=1}^n \sum_{j=1}^m \nabla_\Sigma \gamma_{i,j} u^u_{i,t,j})).
\end{aligned}
$$

The first row of F is identical to that for the test of zero bias (as expected), concatenated by zeros, since $E\partial\varepsilon^c_{1,t,\tau}(\phi)/\partial vech(\Sigma) = 0$. In contrast, $E\partial\varepsilon^c_{1,t,\tau}(\phi) y^c_{1,t,\tau}(\phi)/\partial vech(\Sigma) \neq 0$, and hence the second row of F includes terms contributing to the asymptotic variance due to estimation error in $\hat{\Sigma}_t$.

3.3 Faust-Wright efficiency test

For the purpose of testing the efficiency of the Greenbook forecasts constructed by the staff at the Federal Reserve Board of Governors, Faust and Wright (2008) suggest an alternative formulation of the Mincer-Zarnowitz regression designed to "soak up" the impact of conditioning on the values of variables not in the time t information set. In the notation of our paper, as well as our approach to conditioning, their regression takes the form

$$
\begin{aligned}
\hat{\varepsilon}^c_{1,t,\tau} &= \alpha_0 + \alpha_1(\hat{y}^u_{1,t,\tau} - \hat{y}^c_{1,t,\tau}) + \alpha_2 \hat{y}^c_{1,t,\tau} + error \qquad (1) \\
&= \hat{g}_t'\alpha + error,
\end{aligned}
$$

where $\hat{g}_t = (1, \hat{y}^u_{1,t,\tau} - \hat{y}^c_{1,t,\tau}, \hat{y}^c_{1,t,\tau})'$.

At first blush this regression looks like those discussed in the previous section. And yet there is one major difference that precludes directly applying the results in West and McCracken (1998): α_1 need not be zero under the null of efficiency which, here, is represented by $\alpha_2 = 0$ (and perhaps $\alpha_0 = 0$ if a joint test of efficiency and zero bias is desired). Since the proofs in West and McCracken (1998) explicitly require that $E\hat{\varepsilon}^c_{1,t,\tau}\hat{g}_t = 0$, and hence all elements of α are zero, their results are not applicable.

Regardless, asymptotic normality of the coefficients can be established using the more general results in West (1996) along with the Delta method. To do so, first define the (8×1) function $f_{t,\tau}(\hat{\phi}_t) = (vech_{-1}(\hat{g}_t\hat{g}_t')', \hat{g}_t'\hat{\varepsilon}^c_{1,t,\tau})'$, where the notation $vech_{-1}$ denotes the $vech$ operator but omits the first element (that associated with the intercept). There then exists a

twice continuously differentiable function $q(\cdot)$:$\mathbb{R}^8 \to \mathbb{R}^3$ satisfying $\hat{\alpha} = q(P^{-1}\sum_{t=R}^{T-\tau} f_{t,\tau}(\hat{\phi}_t))$ such that $\alpha_0 = \iota_1' q(Ef_{t,\tau}) = 0$ and $\alpha_2 = \iota_3' q(Ef_{t,\tau}) = 0$ under the null hypothesis. Finally, define $\nabla q(Ef_{t,\tau}) = \partial q(Ef_{t,\tau})/\partial Ef_{t,\tau}$. The following Theorem provides the asymptotic distribution of $P^{1/2}(\hat{\alpha} - \alpha)$.

Theorem 2 *Maintain assumptions A1-A5.* $P^{1/2}(\hat{\alpha}-\alpha) \to^d N(0, V)$ *with* $V = \nabla q(Ef_{t,\tau})'\Omega\nabla q(Ef_{t,\tau})$ *and*

$$\Omega = S_{ff} + \lambda_{fh}(FBS_{fh}' + S_{fh}B'F') + \lambda_{hh}FBS_{hh}B'F',$$

where $S_{ff} = \lim_{P,R\to\infty} Var(P^{-1/2}\sum_{t=R}^{T-\tau} f_{t,\tau})$, $S_{hh} = \lim_{T\to\infty} Var(T^{-1/2}\sum_{s=1}^{T-1} h_{s+1})$, *and* $F = E(\partial f_{t,\tau}(\phi)/\partial\phi)$, *and* $S_{fh} = \lim_{P,R\to\infty} Cov(P^{-1/2}\sum_{t=R}^{T-\tau} f_{t,\tau}, P^{-1/2}\sum_{t=R}^{T-\tau} h_{s+1}')$.

Again we find that the regression coefficients are asymptotically normal. But in contrast to the results in the previous section, only α_0 and α_2 are necessarily zero when the VAR is correctly specified. Also, in contrast to the previous results, there does not appear to be any way of simplifying the formulation of the asymptotic variance V. Both F_β and F_Σ are non-zero and hence estimation error from both $\hat{\beta}_t$ and $\hat{\Sigma}_t$ contribute to the asymptotic variance. In particular, while tedious, the formulas for $F = (F_1', ..., F_8')'$ $(8 \times (\dim vec(\Lambda) + \dim vech(\Sigma)))$ can be derived explicitly and take the following form:

$$\begin{aligned}
F_1 &= (E\sum_{i=1}^n\sum_{j=1}^m \gamma_{i,j}\Theta_{i,t,j}, 0_{n(n+1)/2}) \\
F_2 &= (E(\Theta_{1,t,\tau} - \sum_{i=1}^n\sum_{j=1}^m \gamma_{i,j}\Theta_{i,t,j}), 0_{n(n+1)/2}) \\
F_3 &= (2E(y_{1,t,\tau}^c - y_{1,t,\tau}^u)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\beta\gamma_{i,j}u_{i,t,j}^u), 2E(y_{1,t,\tau}^c - y_{1,t,\tau}^u)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u)) \\
F_4 &= (E(\sum_{i=1}^n\sum_{j=1}^m \gamma_{i,j}y_{1,t,\tau}^u\Theta_{i,t,j}) - 2E(y_{1,t,\tau}^c - y_{1,t,\tau}^u)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\beta\gamma_{i,j}u_{i,t,j}^u), \\
&\quad -2E(y_{1,t,\tau}^c - y_{1,t,\tau}^u)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u)) \\
F_5 &= (2E(y_{1,t,\tau}^c)(-\Theta_{1,t,\tau} + \sum_{i=1}^n\sum_{j=1}^m(-\gamma_{i,j}\Theta_{i,t,j} + \nabla_\beta\gamma_{i,j}u_{i,t,j}^u)), 2E(y_{1,t,\tau}^c)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u)) \\
F_6 &= (E(-\Theta_{1,t,\tau} + \sum_{i=1}^n\sum_{j=1}^m \gamma_{i,j}\Theta_{i,t,j}), 0_{n(n+1)/2}) \\
F_7 &= (-E\sum_{i=1}^n\sum_{j=1}^m \nabla_\beta\gamma_{i,j}u_{i,t,j}^u u_{1,t,\tau}^c + E\sum_{i=1}^n\sum_{j=1}^m \nabla_\beta\gamma_{i,j}u_{i,t,j}^u(y_{1,t,\tau}^c - y_{1,t,\tau}^u), \\
&\quad -E\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u u_{1,t,\tau}^c + E\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u(y_{1,t,\tau}^c - y_{1,t,\tau}^u)) \\
F_8 &= (E(y_{1,t,\tau}^u)(-\Theta_{1,t,\tau} + \sum_{i=1}^n\sum_{j=1}^m \gamma_{i,j}\Theta_{i,t,j}) - E(y_{1,t,\tau}^c - y_{1,t,\tau}^u)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\beta\gamma_{i,j}u_{i,t,j}^u) \\
&\quad +E(u_{1,t,\tau}^c)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\beta\gamma_{i,j}u_{i,t,j}^u), -E(y_{1,t,\tau}^c)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u) \\
&\quad +E(u_{1,t,\tau}^c)(\sum_{i=1}^n\sum_{j=1}^m \nabla_\Sigma\gamma_{i,j}u_{i,t,j}^u)).
\end{aligned}$$

Whereas Faust and Wright (2008) intended their test for use with forecasts such as Greenbook, it could also be used with the VAR-based forecasts considered in this paper.

Their efficiency test is applicable to the minimum-MSE conditional forecasts considered in this paper as well as to the policy shock-based conditional forecasts described in section 2 one might also consider. That said, under the minimum-MSE approach to conditional forecasts, one might expect the Mincer-Zarnowitz regression to have greater power than that proposed by Faust and Wright (2008) because the former imposes the valid restriction $\alpha_1 = 0$ while the latter estimates it freely. Monte Carlo results below general support this intuition.

3.4 Equal accuracy test

In the previous three sections we described tests of predictive ability related to the properties of the conditional forecast error $\hat{\varepsilon}^c_{1,t,\tau}$. In each case, the null hypothesis relates to a property of the forecast error that suggests that the model is useful for constructing conditional forecasts. Rejecting the null hypothesis is indicative of a flaw in the model that might cause future conditional forecasts to be ill-behaved.

One property we have not discussed is the accuracy of the conditional forecast. Intuitively, one would expect that after conditioning on future values of those variables in the hypothetical scenario, the conditional forecast would be more accurate than an unconditional forecast that does not utilize any future information. In other words, if our conditional forecast is good we would expect $E(\varepsilon^c_{1,t,\tau})^2 - E(\varepsilon^u_{1,t,\tau})^2 < 0$. In fact, as also noted in Herbst and Schorfheide (2012), when the forecasts are of the minimum-MSE variety this should be the case.[6]

Unfortunately, this "good" property of the conditional forecast does not give us a workable null hypothesis under which to derive an asymptotic distribution and conduct inference. In fact, given our methodological approach of conditioning on future observations of those variables in a scenario, the null hypothesis $E(\varepsilon^c_{1,t,\tau})^2 - E(\varepsilon^u_{1,t,\tau})^2 = 0$ would require the model to be misspecified. Because of this, one approach to inference would be to derive the asymptotic distribution of the moment condition $P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2)$ assuming a misspecified VAR and reject the null in the lower tail of this distribution. Unfortunately, rejecting the null in the lower tail does not imply that the model is "good" in any sense since the model could still be misspecified and yet $E(\varepsilon^c_{1,t,\tau})^2 - E(\varepsilon^u_{1,t,\tau})^2 < 0$.

We therefore take a different approach to inference, one that continues our strategy of maintaining a correctly specified VAR under the null. Again, under minimum-MSE condi-

[6] This need not hold under the policy shock-based conditioning described in section 2.

tioning, correct specification implies the existence of a non-negative constant k satisfying $E(\varepsilon^c_{1,t,\tau})^2 - E(\varepsilon^u_{1,t,\tau})^2 = k$. This constant depends on the VAR regression parameters A_j and residual variance Σ and can be derived explicitly as a function of ϕ. To do so, first define $\Psi_j\Sigma^{1/2}$ as the matrix of orthogonalized impulse responses after j periods and let

$$D = \begin{pmatrix} \Sigma^{1/2} & 0 & & 0 & 0 \\ \Psi_1\Sigma^{1/2} & \Sigma^{1/2} & & 0 & 0 \\ & & \dots & \Sigma^{1/2} & 0 \\ \Psi_{\max(\tau,m)}\Sigma^{1/2} & \Psi_{\max(\tau,m)-1}\Sigma^{1/2} & & \Psi_1\Sigma^{1/2} & \Sigma^{1/2} \end{pmatrix}.$$

Now let $\tilde{D}$ denote the matrix formed by those rows in D associated with a conditioning restriction.[7] For example, if $n = 2$ and we condition on future values of the second element of the VAR at both the first and second horizons ($m = 2$), $\tilde{D}$ consists of the (2×4) matrix formed by stacking the second and fourth rows of D. Straightforward algebra then implies $k(\phi) = \iota_1' D\tilde{D}'(\tilde{D}\tilde{D}')^{-1}\tilde{D}D'\iota_1$.

With this constant $k(\phi)$ in hand we consider testing for "equal accuracy" (below, we will just generally refer to this as testing MSE accuracy) using the appropriately re-centered Diebold and Mariano (1995) and West (1996)-type test of predictive ability:

$$P^{1/2}\hat{\alpha} = P^{-1/2}\sum\nolimits_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2 - k(\hat{\phi}_T)).$$

Theorem 3 *Maintain assumptions A1-A5.* $P^{1/2}\hat{\alpha} \to^d N(0,\Omega)$ *with*

$$\begin{aligned} \Omega &= S_{ff} + 2\lambda_{fh}FBS'_{fh} - 2\frac{\pi}{1+\pi}KBS'_{fh} - 2\frac{\pi}{1+\pi}FBS_{hh}B'K' \\ &\quad + \lambda_{hh}FBS_{hh}B'F' + \frac{\pi}{1+\pi}KBS_{hh}B'K', \end{aligned}$$

where $S_{fh} = \lim_{P,R\to\infty} Cov(P^{-1/2}\sum_{t=R}^{T-\tau}((\varepsilon^c_{1,t,\tau})^2 - (\varepsilon^u_{1,t,\tau})^2 - k), P^{-1/2}\sum_{t=R}^{T-\tau} h'_{s+1})$, $S_{ff} = \lim_{P,R\to\infty} Var(P^{-1/2}\sum_{t=R}^{T-\tau}((\varepsilon^c_{1,t,\tau})^2 - (\varepsilon^u_{1,t,\tau})^2 - k))$, $S_{hh} = \lim_{T\to\infty} Var(T^{-1/2}\sum_{s=1}^{T-1} h_{s+1})$, $F = E(\partial((\varepsilon^c_{1,t,\tau}(\phi))^2 - (\varepsilon^u_{1,t,\tau}(\phi))^2)/\partial\phi)$, *and* $K = \partial k(\phi)/\partial\phi$.

Theorem 3 implies that an appropriately centered DM/W test of equal predictive ability is asymptotically normal. While this is precisely what is shown in DM/W, there is an important distinction: estimation error affects the asymptotic distribution. This contrasts with the results in West (1996). In that paper it is shown that when the forecasts associated with two non-nested OLS estimated linear models are evaluated under quadratic loss, $F = 0$ (and for that matter, $K = 0$) and hence estimation error is asymptotically irrelevant.

[7] This notation is taken directly from Jarocinski (2010) except we substitute D for R.

Mathematically, the key feature driving this difference is the presence of the weights $\gamma_{i,j}$ as well as the centering constant k, and, in particular the fact that they are functions of ϕ. Let $\nabla_\beta \gamma_{i,j} = \partial\gamma_{i,j}(\phi)/\partial\beta$ and $\nabla_\Sigma \gamma_{i,j} = \partial\gamma_{i,j}(\phi)/\partial vech(\Sigma)$. Straightforward algebra reveals that $F = (F_\beta, F_\Sigma)$, where

$$F_\beta = -\sum_{i=1}^{n}\sum_{j=1}^{m} \nabla_\beta \gamma_{i,j} E u^u_{1,t,\tau} u^u_{i,t,j} + 2E(\sum_{i=1}^{n}\sum_{j=1}^{m} \gamma_{i,j} u^u_{i,t,j})(\sum_{i=1}^{n}\sum_{j=1}^{m} \nabla_\beta \gamma_{i,j} u^u_{i,t,j})$$

and

$$F_\Sigma = -\sum_{i=1}^{n}\sum_{j=1}^{m} \nabla_\Sigma \gamma_{i,j} E u^u_{1,t,\tau} u^u_{i,t,j} + 2E(\sum_{i=1}^{n}\sum_{j=1}^{m} \gamma_{i,j} u^u_{i,t,j})(\sum_{i=1}^{n}\sum_{j=1}^{m} \nabla_\Sigma \gamma_{i,j} u^u_{i,t,j}).$$

Since F_β and F_Σ (as well as unreported K_β and K_Σ) are both non-zero, estimation error in both $\hat{\beta}_t$ and $\hat{\Sigma}_t$ affect the asymptotic distribution through the asymptotic variance.

3.5 A bootstrap approach to inference

For each test of bias and efficiency we are able to establish that $P^{1/2}(\hat{\alpha} - \alpha)$ is asymptotically normal. We are also able to establish asymptotic normality of the test of MSE accuracy. As such, one approach to inference is to construct a consistent estimate of the asymptotic variance and then compare the standardized statistic $P^{1/2}(\hat{\alpha}_j - \alpha_j)/\hat{\Omega}_j^{1/2}$ to standard normal critical values. As discussed in West (1996) as well as West and McCracken (1998), many of the components of the asymptotic variance are easily estimated, while other elements can be more complicated. Among the easiest components are $\hat{\pi} = P/R$, $\hat{B}_\beta = ((T^{-1}\sum_{s=1}^{T-1} x_s x_s')^{-1} \otimes I_n)$, and $B_\Sigma = I_{n(n+1)/2}$. In addition, standard HAC estimators can be used to construct estimates of S_{ff}, S_{fh}, and S_{hh} given $f_{t,\tau}(\hat{\phi}_t)$ and $\hat{h}_{t+1} = (vec(\hat{\varepsilon}_{t+1} x_t')', vech(\hat{\varepsilon}_{t+1}\hat{\varepsilon}_{t+1}' - \hat{\Sigma}_T)')'$, where $\hat{\varepsilon}_{t+1}$ denotes the residuals from the OLS-estimated VAR.

Unfortunately, estimating F, $\nabla q(Ef_{t,\tau})$, and K is significantly more difficult even when, for example, formulas for F are provided as they are in the text. As a result, we suggest and use in our applications a bootstrap approach to inference. However, it's important to note that the choice of bootstrap is non-trivial. Recall that for many of our results the asymptotic distribution is affected by estimation error in both $\hat{\beta}_t$ and $\hat{\Sigma}_t$. This implies that most bootstrap procedures, including that of Goncalves and Kilian (2004) which captures the asymptotic behavior of $\hat{\beta}_t$ but not $\hat{\Sigma}_t$, are not generally applicable for the testing procedures in this paper. We therefore use a VAR-based, moving-block residual based bootstrap developed in Bruggemann, Jentsch, and Trenkler (2014), which is specifically designed to

capture the asymptotic behavior in $\hat{\beta}_t$ and $\hat{\Sigma}_t$ simultaneously (along with conditional heteroskedasticity in the VAR's innovations). To implement their bootstrap in our setting, we estimate the VAR of interest with the full sample of data and save the residuals and coefficients. We then bootstrap the residuals using a moving block bootstrap of 40 observations.[8] We use these residuals and the autoregressive structure of the VAR to obtain an artificial time series y_t^*.[9] In each bootstrap replication, the bootstrapped data are used to estimate the VAR forecasting model at each forecast origin and generate artificial, out-of-sample forecasts. These forecasts and associated forecast errors are used to produce bias, efficiency, and MSE accuracy test statistics. Critical values are simply computed as percentiles of the bootstrapped test statistics.

4 Monte Carlo Analysis

This section presents a Monte Carlo analysis of the finite-sample properties of tests for bias, efficiency, and MSE accuracy applied to unconditional and conditional forecasts from VAR models. In all cases, consistent with the preceding analytics, we produce and evaluate forecasts conditioned on a path for a pseudo-policy variable that is the actual future path. Using the simulated data, we conduct forecast inference based on both normal distributions with standard errors that abstract from parameter estimation error and bootstrapped test distributions.

In these experiments, we use bivariate and trivariate VARs as data-generating processes (DGPs). In the reported results, we form (iterated multi-step) forecasts using OLS estimates of VARs. We focus on forecasts computed under a recursive (expanding window) estimation scheme, but we provide some results for a rolling window estimation scheme. For the conditional forecasts, we concentrate on the minimum-MSE approach to conditional forecasting used in the theoretical analysis above.

In all simulations, based on 2000 Monte Carlo draws, we report the percentage of Monte Carlo trials in which the null of no bias, efficiency, or MSE accuracy is rejected — the percentage of trials in which the sample test statistics fall outside (two-sided) critical values. In the reported results, the tests are compared against 10% critical values. Using 5% critical values yields similar findings.

[8] In our Monte Carlo analysis, before settling on a block size of 40, we considered a grid of sizes up through 50, and found differences across block choices to be fairly small.

[9] The initial observations are selected by sampling from the actual data as in Stine (1987).

We proceed by first detailing the data-generating processes and other aspects of experiment design and then presenting the results.

4.1 Monte Carlo design

For each DGP, we generate data using draws of innovations and the autoregressive structure of the DGP. The initial observations necessitated by the lag structure of each DGP are generated with draws from the unconditional normal distribution implied by the DGP. With quarterly data in mind, we report results for forecast horizons of 1, 2, and in some cases, 4 periods. We consider sample sizes of $R, P =$ 50,100; 50,150; 100,50; and 100,100. We use DGPs 1, 1G, and 2 to evaluate size properties and DGPs 3 and 4 to evaluate power.

DGP 1 is a bivariate VAR(1), with regression coefficients given in the first panel of Table 1 and an error variance-covariance matrix of:

$$\text{var}(\varepsilon_t) = \begin{pmatrix} 1.0 & \\ 0.5 & 1.0 \end{pmatrix}.$$

In order to assess test reliability with conditionally heteroskedastic forecast errors, we also consider a version of DGP 1 with GARCH innovations, denoted DGP 1G. The VAR has the same regression coefficients as given for DGP 1 in the first panel of Table 1. In the GARCH structure, taken from an experiment design in Bruggemann, Jentsch, and Trenkler (2014), let e_t denote a 2×1 innovation vector that is distributed $N(0, I_2)$. Let $v_{i,t} = \sigma_{i,t} e_{i,t}$, where $\sigma^2_{i,t} = 0.2 + 0.05 v^2_{i,t-1} + 0.75 \sigma^2_{i,t-1}$, such that $v_{1,t}$ and $v_{2,t}$ are independent GARCH(1,1) processes. The VAR innovations are constructed as Sv_t, where S is

$$\text{var}(\varepsilon_t) = \begin{pmatrix} 1.0 & 0.0 \\ 0.5 & 0.866 \end{pmatrix}.$$

With this formulation, the unconditional error variance matrix is the same as in the homoskedastic version of DGP 1.

DGP 2 is a trivariate VAR(2), with regression coefficients given in the second panel of Table 1 and an error variance-covariance matrix of:

$$\text{var}(\varepsilon_t) = \begin{pmatrix} 9.265 & & \\ 0.296 & 1.746 & \\ 0.553 & 0.184 & 0.752 \end{pmatrix}.$$

We set the parameters of DGP 2 to equal OLS estimates of a VAR in GDP growth, inflation less a survey-based measure of trend inflation (see, e.g., Clark and Doh (2014), Faust and Wright (2013)), and the federal funds rate less a survey-based measure of trend, over a sample of 1961-2007.

To evaluate power properties, DGPs 3 and 4 impose some parameter breaks on the specification of the bivariate DGP 1. In DGP 3, the breaks consist of one-time shifts in the intercept of the $y_{1,t}$ equation and the slope and intercept coefficients of the $y_{2,t}$ equation. The pre- and post-break coefficients are given in the last panel of Table 1; the error variance-covariance matrix is kept constant, at the setting used with DGP 1 (see above). The break is imposed to occur at period $R+1$, the date of the first out-of-sample forecast. In DGP 4, the single break takes the form of a shift in the error correlation ρ, from 0.5 to -0.5, at period $R+1$. In this DGP, the other VAR parameters are stable over time, at the DGP 1 values given in Table 1.

In experiments with DGPs 1, 1G, 3, and 4, to keep the conditioning relatively simple and speed the Monte Carlo replications, we report results for forecast horizons of 1 and 2 periods ahead. In these cases, we forecast the first variable of the VAR in periods $t+1$ and $t+2$ (i.e., $y_{1,t+1}$ and $y_{1,t+2}$), conditional on the actual values of the second variable in those periods (i.e., $y_{2,t+1}$ and $y_{2,t+2}$). In experiments with DGP 2, we extend the conditioning horizon to 4 quarters ahead. In this case, we forecast variables $y_{1,t+\tau}$ and $y_{2,t+\tau}$ for $\tau = 1,\ldots,4$ conditional on $y_{3,t+\tau}$ equaling its actual values in periods $t+1$ through $t+4$ — that is, conditional on $\hat{y}^c_{3,t,\tau} = y_{3,t+\tau}, \tau = 1,\ldots,4$.

4.2 Test implementation

For convenience, we list here the regressions and test statistics we use, referring to variable 1 of the VAR for convenience (in DGP 2, we consider forecasts of an additional variable):

$$\hat{\varepsilon}^i_{1,t,\tau} = \alpha_0 + e_{t,\tau},\ i = u,c,\ \text{bias: } \alpha_0\ t\text{-stat} \tag{2}$$

$$\hat{\varepsilon}^i_{1,t,\tau} = \alpha_0 + \alpha_1 \hat{y}^i_{1,t,\tau} + e_{t,\tau},\ i = u,c,\ \text{M-Z: } \alpha_1\ t\text{-stat} \tag{3}$$

$$\hat{\varepsilon}^c_{1,t,\tau} = \alpha_0 + \alpha_1(\hat{y}^u_{1,t,\tau} - \hat{y}^c_{1,t,\tau}) + \alpha_2 \hat{y}^c_{1,t,\tau} + e_{t,\tau},\ \text{F-W: } \alpha_2\ t\text{-stat} \tag{4}$$

$$(\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2 = \alpha_0 + e_{t,\tau},\ \text{MSE accuracy: } \alpha_0 = k(\hat{\phi}_T)\ t\text{-stat} \tag{5}$$

To test bias, we regress (equation (2)) forecast errors (either unconditional or conditional) on a constant and form the t-statistic for the null of a coefficient of zero. To test efficiency, for all types of forecasts, we consider the Mincer-Zarnowitz efficiency test of equation (3). For conditional forecasts, we also consider the t-statistic of the coefficient on the conditional forecast in the Faust and Wright (2008) efficiency regression (4). Finally, we use the regression (5) to test whether the difference in MSEs for the unconditional and

conditional forecasts is different from the population level-value implied by the VAR's parameters (where, as noted above, that population value is a function of the VAR's slope coefficients and error variance-covariance matrix).

In all cases, the forecast error for horizon τ should follow an MA($\tau - 1$) process. At the 1-step horizon, we form test statistics using the simple OLS estimate of the variance. At longer horizons, we use a rectangular kernel, with $\tau - 1$ lags.[10]

4.3 Results

Tables 2 through 5 provide size results from Monte Carlo experiments with DGPs 1, 1G, and 2, in which the conditional forecasts are computed with the minimum-MSE approach.

These results indicate that comparing tests of bias, efficiency, and MSE accuracy against standard normal critical values will often be unreliable. More specifically:

- Under a recursive scheme, without correction for the effects of parameter estimation error, tests of bias — both unconditional and conditional — range from being about correctly sized (DGP 1) to modestly oversized (DGP 2). Oversizing is somewhat more likely at longer forecast horizons, probably because of difficulties with the finite-sample precision of HAC variance estimates.

- Under a recursive scheme, tests of efficiency range from being almost correctly sized (DGP 1, 1-step horizon) to significantly oversized (DGP 3). Size is comparable for the M-Z and F-W tests applied to conditional forecasts and the M-Z test applied to unconditional forecasts. The size distortions generally rise as the forecast horizon increases, likely reflecting imprecision in the HAC variance estimate. For example, for forecasts of $y_{2,t}$ in experiments with DGP 2 and $R, P = 100{,}100$, the rejection rate for the normal-based M-Z efficiency test applied to conditional forecasts rises from 23.0 percent at the 1-step horizon to 29.0 percent at the 2-step horizon and 36.1 percent at the 4-step horizon. Although not shown in the tables, the oversizing seems to stem partly from a left-ward shift of the finite-sample distribution of the test statistics. This shift will become more evident in the application results of the next section.

- With recursively produced forecasts, the MSE accuracy test often, although not always, leads to significant undersizing. For example, with DGP 2, $R, P = 50{,}100$,

[10] In the rare occasions in which the result variance estimator is not positive semi-definite, we instead use the HAC estimator of Newey and West (1987), with $2(\tau - 1)$ lags.

and the 1-step forecast horizon, the rejection rate for the MSE test is 3.5 percent for variable 1 and 14.3 percent for variable 2.

- Under the rolling scheme, the failure to correct standard errors for the effects of parameter estimation error associated with forecast model estimation can have large consequences. In particular, rejection rates for bias tests are materially lower under the rolling scheme than the recursive, such that bias tests applied to rolling forecasts are often modestly or significantly undersized. In the experiments in which $P/R > 1$, rejection rates for efficiency tests are higher under the rolling scheme than the recursive, such that efficiency tests applied to rolling forecasts are often significantly oversized. Finally, tests of MSE accuracy in DGP 1 applied under the rolling scheme range from significantly to slightly undersized.

- Under the recursive scheme, conditional heteroskedasticity in the data-generating process does not materially affect the results described above. Broadly, size results under DGP 1G (Table 5) are fairly similar to those under DGP 1 (Table 2). The oversizing of efficiency tests is a little greater in the GARCH results than in the conditionally homoskedastic results, but otherwise results are fairly similar.

In comparison, conducting inference on the basis of our proposed bootstrap yields more reliable tests of bias, efficiency, and accuracy. More specifically:

- For both unconditional and conditional forecasts, the median rejection rate across all bias tests in Tables 2-5 is about 11.4 percent, with a minimum of about 9 percent and maximum of about 13-14 percent.

- For each type of efficiency test, the median rejection rate across all experiments in Tables 2-5 is a little less than 10 percent. In some settings, however, the tests can be either slightly-to-modestly undersized or slightly-to-modestly oversized. For example, applied to unconditional forecasts the M-Z test of efficiency has a minimum of 4.6 percent and maximum of 14.1 percent across experiments. Applied to conditional forecasts the M-Z test's rejection rate ranges from 5.9 percent to 14.3 percent.

- Size performance is more variable for the MSE test (unconditional versus minimum-MSE conditional) than the bias and efficiency tests. On average, the bootstrap is reasonably reliable: the median rejection rate across all MSE tests in Tables 2-5 is

8.9 percent. But the rejection rate ranges from 3.2 to 17.1 percent, depending on the experiment, variable, horizon, etc.

- With bootstrap inference, using a rolling scheme for estimation or a DGP with GARCH yields results similar to those obtained for recursive estimation of a DGP with conditionally homoskedastic innovations.

Tables 6 and 7 provide power results from Monte Carlo experiments with DGPs 3 and 4, in which the conditional forecasts are computed with the minimum-MSE approach and the forecasting scheme is recursive. We focus on power results obtained under bootstrap critical values, because using normal critical values (without correction for parameter estimation error) does not yield accurately sized tests. These experiments yield the following findings for power.

- With a break in VAR coefficients (DGP 3), tests for bias have relatively good power. For example, with R, P = 100,100, the 2-step ahead rejection rate is 85.0 percent for the unconditional forecast and 74.8 percent for the conditional forecast. As this example indicates, the bias test for unconditional forecasts has somewhat more power than the bias test for conditional forecasts, likely because conditioning on the actual future path of $y_{2,t}$ reduces shifts in $y_{2,t}$ as a potential source of instability in the forecast of $y_{1,t}$.

- In the same DGP, tests for efficiency have at best modest power, with the M-Z test applied to unconditional forecasts having a little more power than the M-Z or F-W tests applied to conditional forecasts. In the same example (R, P = 100,100 and a 2-step forecast horizon), the rejection rate is 17.3, 7.9, and 7.1 percent for the M-Z unconditional, M-Z conditional, and F-W conditional tests.

- Similarly, in the DGP with a break in VAR coefficients, the test of MSE accuracy has at least a little power. In the same example, the power of the MSE accuracy test is 14.9 percent (Table 6).

- In the DGP with a break in error correlations (Σ) instead of the VAR's slope coefficients, the power rankings of the tests are essentially reversed. In DGP 4, the MSE accuracy test has the best power, ranging from about 40 percent (with R, P = 50,150) to as much as 94.5 percent (R, P = 100,50, 1-step ahead forecast horizon). The M-Z

efficiency test applied to conditional forecasts ranks second in power, yielding a rejection rate as high as 75.7 percent (R, P = 100,50, 2-step ahead forecast horizon), but also as low as 11.9 percent (with R, P = 50,150, 1-step ahead forecast horizon). Neither the F-W test for the efficiency of conditional forecasts nor the M-Z test for the efficiency of unconditional forecasts have much power. Bias tests also lack power in the face of a shift in the error correlation (in Σ), likely due to the absence of any mean shifts that would lead to bias in forecasts.

In summary, the Monte Carlo results are generally consistent with the analytical findings described above. With conditional forecasts, comparing tests of bias, efficiency, and MSE accuracy against standard normal critical values does not yield generally reliable results. A likely reason is that, in many cases, the asymptotic distributions of the tests are affected by parameter estimation error (from forecast model estimation). While the statistics of interest are normally distributed, parameter estimation error affects the appropriate variance in a way that is often very complicated. Our proposed (and easily implemented) bootstrap procedure yields tests with reasonably reliable size properties. As to power, our results show that shifts in error correlations will lead to rejections of MSE accuracy (unconditional versus conditional) and the efficiency of conditional forecasts, but not rejections of the efficiency of unconditional forecasts.

5 Empirical Application

In our empirical application, we examine forecasts obtained with a VAR (estimated by OLS) in the variables included in the DSGE model of Smets and Wouters (2007). The Smets-Wouters model and variants of it are widely used in other research and practice (e.g., Del Negro, Giannoni, and Schorfheide (2015)), and a number of studies, including Smets and Wouters (2007), Edge and Gurkaynak (2010), and Gurkaynak, Kisacikoglu, and Rossi (2013), have compared the accuracy of forecasts from the Smets-Wouters model to forecasts from VARs in the same variables.

This section proceeds by first providing data and forecasting details and then providing results. In the interest of brevity, we report forecast results for a selected subset of variables: growth in GDP, growth in private fixed investment, and inflation. We do not report forecasts for the federal funds rate because we are interested in forecasts of the other variables conditional on particular paths of the federal funds rate.

Variables in VAR of Application

GDP, chain-weighted	$400\Delta \ln \text{GDP/population}$
Consumption (PCE), chain-weighted	$400\Delta \ln \text{PCE/population}$
Private fixed investment (PFI), chain-weighted	$400\Delta \ln \text{PFI/population}$
Aggregate hours worked (hours), non-farm business sector	$400\Delta \ln \text{hours/population}$
Real hourly compensation (comp), non-farm business sector	$400\Delta \ln \text{comp/GDPPI}$
Inflation, GDP price index (GDPPI)	$400\Delta \ln \text{GDPPI}$
Federal funds rate	level

5.1 Data and Sample

The model variables consist of GDP, consumption, private fixed investment, real hourly labor compensation, hours worked, inflation in the GDP price index, and the federal funds rate, all measured on a quarterly basis.[11] The GDP, consumption, investment, and hours worked variables are defined in per-capita terms, and these variables plus real compensation enter the model as annualized growth rates. To transform to per-capita terms, we normalize by the non-institutional population, aged 16 and over. We obtained all data from the FAME database of the Federal Reserve Board, except for the population series, which we obtained from the (publicly available) Federal Reserve Board's FRB/US model database.[12] The table below lists the variables and their transformations.

We generate and evaluate forecasts for a sample of 1991 through 2007, at horizons of 1, 2, 4, and 8 quarters. We start forecasting in 1991 and not sooner because, through the 1980s, inflation was still trending down, presumably due to a deliberate disinflation effort by the Federal Reserve. We stop our forecast evaluation in 2007:Q4 to avoid possible complications of the zero lower bound constraints that became relevant in subsequent years.

In all cases, we estimate the model with a sample that starts in 1961:Q1. We begin by estimating with data from 1961:Q1 through 1990:Q4 and forming forecasts for 1991:Q1 through 1992:Q4. We then proceed by moving to a forecast origin of 1991:Q2, estimating the model with data from 1961:Q1 through 1991:Q1 and forming forecasts for 1991:Q2 through 1993:Q1. We proceed similarly through time (under the recursive scheme), up through 2007:Q3, to obtain a sample of forecasts from 1991 through 2007. At each point in time, we produce conditional forecasts based on the actual future path of the federal funds

[11] For various reasons, our measurements of these variables differ some from those of Smets and Wouters (2007) and other studies of similar models. For example, to be consistent with the definitions of the NIPA accounts, we use real consumption defined as nominal divided by the consumption deflator, whereas Smets and Wouters used nominal consumption divided by the GDP deflator.

[12] This series is adjusted for breaks created by changes in population controls that can occur at the beginning of each year.

rate over the 8-step forecast horizon. In the presented results, in the interest of brevity, we focus on conditional forecasts produced with the minimum-MSE approach.

5.2 Results

Tables 8 and 9 provide the results of bias, efficiency, and MSE accuracy tests applied to growth, investment, and inflation forecasts. In particular, Table 8 reports test statistics (t-statistics) along with bootstrapped 10 percent critical values. Table 9 reports the MSEs of the unconditional and conditional forecasts, along with the t-test for MSE accuracy and its 10 percent critical values.

Broadly, using the bootstrap critical values that our Monte Carlo analysis show to be more reliable than standard normal critical values, Table 8 provides consistent evidence of bias in forecasts of GDP growth and inefficiency in forecasts of inflation. In the case of GDP growth forecasts, both unconditional and conditional forecasts look to be biased at all or most horizons considered. For inflation, there is somewhat more evidence of inefficiency in the conditional forecasts than the unconditional forecasts. For the unconditional forecast of inflation, the Mincer-Zarnowitz test rejects efficiency at horizons of one and two steps but not at longer horizons. For the conditional forecasts of inflation, our proposed test rejects efficiency at all horizons, and the Faust-Wright test rejects at horizons of one through four quarters. In addition, there is some more limited evidence of inefficiency in forecasts of GDP growth (both unconditional and conditional, at just the two-step horizon) and investment growth (conditional at horizons of two and four quarters). Overall, inefficiency is rejected a little more often for conditional forecasts than unconditional forecasts.

Table 8 also shows that using standard normal critical values would overstate the evidence of forecast inefficiency, as occurs in some of the Monte Carlo experiments of the previous section. There are a number of cases (e.g., M-Z tests of efficiency of unconditional forecasts of investment growth) in which comparing the reported test statistics against normal critical values would lead one to reject the null of efficiency but comparing against the more accurate bootstrap critical values does not yield rejection. The reported bootstrap critical values suggest the finite sample distributions of the efficiency test statistics to be shifted to the left of a distribution centered about 0. While not detailed in the Monte Carlo results, in unreported checks we found the same shift present in simulations of DGP 2, which shows significant size distortions of efficiency tests compared against standard normal critical values.

The results in Table 9 indicate that the differences in MSE accuracy between the unconditional and minimum-MSE conditional forecasts are generally not significantly different from what full sample estimates of the VAR imply should obtain (with the exception of one-step ahead forecasts of GDP growth). To be sure, there are differences in the actual MSEs that are somewhat different from the MSEs implied by the VAR's parameters, but these differentials are not often statistically significant. Consider, for example, 4-quarter ahead forecasts of GDP growth. In the out-of-sample forecasts, the difference in MSEs (unconditional less minimum-MSE conditional) is 1.454. The VAR, with full-sample parameter estimates, implies the difference is MSEs could be expected (in population) to be 2.381. However, the t-statistic for the equality of the sample difference and population-implied difference is just -1.490, well above the lower tail of the bootstrap distribution.

Overall, these results suggest some misspecification of the VAR fit to the Smets and Wouters (2007) variable set and used to produce the forecasts of GDP growth, investment, and inflation — both unconditional forecasts and forecasts conditioned on the actual path of the federal funds rate over a two year horizon. Although that misspecification doesn't lead to systematic bias in the forecasts of investment or inflation or cause the difference in the accuracy of unconditional and conditional forecasts to be significantly different than the VAR's parameters imply, it leads to a number of rejections of the efficiency of both unconditional and conditional forecasts.

6 Conclusions

Motivated by the common use of conditional forecasting in both practical forecasting and research, we develop and apply methods for the evaluation of conditional forecasts from VARs, using tests of bias, efficiency, and the MSE accuracy of conditional versus unconditional forecasts. More specifically, we provide analytical, Monte Carlo, and empirical evidence on tests of predictive ability for conditional forecasts from estimated models. Throughout, our intention is to provide usable metrics for evaluating conditional forecasts, in a general sense and in comparison to the accuracy of unconditional forecasts. To do so, we focus on particular forms of conditional forecasts for which interpretation of various null and alternative hypotheses is most straight-forward. In particular, in our analysis, we consider forecasts conditioned on actual future information on some variables in the VAR model.

For these tests, we establish asymptotic normality in the context of VAR-based con-

ditional forecasts. Our results follow from an application of West (1996) and West and McCracken (1998) and as such establish the role of estimation error on the asymptotic distribution. As a practical matter, the standard errors can be quite complex and as such we suggest and consider a bootstrap approach to inference that is valid when even estimation error contributes to the asymptotic variance of the test statistic. Monte Carlo evidence suggests that the tests can be reasonably well sized in samples of the size often seen in macroeconomic applications.

Building on these results, in our application, we evaluate unconditional and conditional forecasts from a common macroeconomic VAR in the Smets and Wouters (2007) set of variables. We produce unconditional and conditional forecasts of GDP growth, investment growth, and inflation, in which the conditional forecasts are based on the actual path of the short-term interest rate over an eight quarter forecast horizon. Using bootstrap critical values, we find evidence of model misspecification in the form of rejections of the efficiency of unconditional and conditional forecasts.

7 Appendix: Theory Details

In this section we provide proofs of the Theorems described in the text. In addition to the notation from Section 2, let $z_{t+r} = (x'_{t+r}, ..., x'_t)'$ where $r = \max(\tau, m)$. Throughout we ignore the finite sample difference between P and $P - \tau + 1$.

Proof of Theorem 1: Given A1 - A5, the proof follows from Theorems 4.1 and 4.2 of West and McCracken (2008).

Proof of Theorem 2: Recall that $f_{t,\tau}(\hat{\phi}_t) = (vech_{-1}(\hat{g}_t\hat{g}'_t)', \hat{g}'_t\hat{\varepsilon}^c_{1,t,\tau})'$. Given Theorem 4.1 of West (1996), the result will follow if A1 - A5 are sufficient for Assumptions 1-4 (W1-W4) in West (1996) when applied to $f_{t,\tau}(\hat{\phi})$ and h_s. W2 and W4 follow immediately from A2 and A5. For W1 and W3 it's useful to recall that every element of $f_{t,\tau}(\hat{\phi}_t)$, $h_s(\phi)$, $\partial f_{t,\tau}(\hat{\phi}_t)/\partial\phi$, $\partial h_s(\phi)/\partial\phi$, $\partial^2 f_{t,\tau}(\hat{\phi}_t)/\partial\phi\partial\phi'$, and $\partial^2 h_s(\phi)/\partial\phi\partial\phi'$ are twice continuously differentiable functions of polynomials of ϕ and are quadratics of z_{t+r}. This follows from (i) the fact that the unconditional forecasts are iterated multistep forecasts from a VAR, (ii) the definition of minimum-MSE conditioning, and (iii) the definition of $f_{t,\tau}(\hat{\phi}_t)$ and $h_s(\phi)$. As such, A1 and A3 suffice for W1. Finally, since (i) fourth order stationarity of w_t implies covariance stationarity of $f_{t,\tau}$, $\partial f_{t,\tau}/\partial\phi$ and h_s, (ii) mixing is preserved by finite dimensioned functions of mixing variables, and (iii) the existence of $8d$ moments for w_t implies the existence of $4d$ moments for $f_{t,\tau}$, $\partial f_{t,\tau}/\partial\phi$ and h_s we find that A1 and A4 suffice for W3 and the proof is complete.

Proof of Theorem 3: Define $f_{t,\tau}(\hat{\phi}_t)$ by rewriting the moment condition as $(\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2 - k(\hat{\phi}_T) = f_{t,\tau}(\hat{\phi}_t) - (k(\hat{\phi}_T) - k)$. We provide the proof separately for the case when $\pi = 0$ and when $\pi > 0$.

(a) Let $\pi = 0$. Assumptions A2 and A4 imply $T^{1/2}(\hat{\phi}_T - \phi) \to^d N(0, BS_{hh}B')$. Since $k(\phi)$ is continuously differentiable in ϕ, the Delta method implies $T^{1/2}(k(\hat{\phi}_T) - k) \to^d N(0, KBS_{hh}B'K')$. Since $\pi = 0$ implies $T/P = o(1)$, $P^{1/2}(k(\hat{\phi}_T) - k) = o_p(1)$ and hence $P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2 - k(\hat{\phi}_T)) = P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2 - k) + o_p(1)$. The result will follow if A1 - A5 are sufficient for assumptions W1-W4 in West (1996) when applied to $f_{t,\tau}(\hat{\phi})$ and h_s. Since the arguments used in the proof of Theorem 2 are applicable when $f_{t,\tau}(\hat{\phi}_t) = (\hat{\varepsilon}^c_{1,t,\tau})^2 - (\hat{\varepsilon}^u_{1,t,\tau})^2 - k$, the result follows from Theorem 4.1 of West (1996).

(b) Let $\pi > 0$. From the proof of (a) it's clear that $P^{1/2}(k(\hat{\phi}_T) - k) = KB(P^{1/2}H(T)) + o_p(1) \to^d N(0, \frac{\pi}{1+\pi}KBS_{hh}B'K')$. In addition, Lemma 4.1 and Theorem 4.1 of West (1996)

imply

$$\begin{aligned} P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2-(\hat{\varepsilon}^u_{1,t,\tau})^2-k) &= P^{-1/2}\sum_{t=R}^{T-\tau} f_{t,\tau}+FB(P^{-1/2}\sum_{t=R}^{T-\tau}H(t))+o_p(1) \\ &\rightarrow {}^dN(0,S_{ff}+2\lambda_{fh}FBS'_{fh}+\lambda_{hh}FBS_{hh}B'F'). \end{aligned}$$

Together these imply

$$\begin{aligned} P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2-(\hat{\varepsilon}^u_{1,t,\tau})^2-k(\hat{\phi}_T)) &= P^{-1/2}\sum_{t=R}^{T-\tau} f_{t,\tau}+FB(P^{-1/2}\sum_{t=R}^{T-\tau}H(t)) \\ -KB(P^{1/2}H(T))+o_p(1) &\rightarrow {}^dN(0,\Omega_0) \end{aligned}$$

for some asymptotic variance Ω_0. The result will follow if we can establish that $\Omega_0=\Omega$ as described in the text. To do this note that we can rewrite $P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2-(\hat{\varepsilon}^u_{1,t,\tau})^2-k(\hat{\phi}_T))$ as

$$P^{-1/2}\sum_{t=R}^{T-\tau}((\hat{\varepsilon}^c_{1,t,\tau})^2-(\hat{\varepsilon}^u_{1,t,\tau})^2-k(\hat{\phi}_T))=[I,FB,-KB]\begin{pmatrix} P^{-1/2}\sum_{t=R}^{T-\tau} f_{t,\tau} \\ P^{-1/2}\sum_{t=R}^{T-\tau}H(t) \\ P^{1/2}H(T)\end{pmatrix}+o_p(1)$$

and hence

$$\Omega_0=[I,FB,-KB]\lim_{P,R\rightarrow\infty}Var(P^{-1/2}\sum_{t=R}^{T-\tau}f_{t,\tau},P^{-1/2}\sum_{t=R}^{T-\tau}H(t)',P^{1/2}H(T)')[I,FB,-KB]'.$$

Straightforward algebra reveals that $\Omega_0=\Omega$ if (i) $\lim_{P,R\rightarrow\infty}Cov(P^{-1/2}\sum_{t=R}^{T-\tau}f_{t,\tau},P^{1/2}H(T)')=\frac{\pi}{1+\pi}S'_{fh}$ and (ii) $\lim_{P,R\rightarrow\infty}Cov(P^{-1/2}\sum_{t=R}^{T-\tau}H(t),P^{1/2}H(T)')=\frac{\pi}{1+\pi}S_{hh}$. We establish both of these below accounting for whether the rolling or recursive schemes are used.

(b-i) Note that $P^{1/2}H(T)=\frac{(PR)^{1/2}}{T}(R^{-1/2}\sum_{s=1}^{R-1}h_{s+1})+\frac{P}{T}(P^{-1/2}\sum_{s=R}^{T-1}h_{s+1})$. Given A4, it is straightforward to show that $\lim_{P,R\rightarrow\infty}Cov(P^{-1/2}\sum_{t=R}^{T-\tau}f_{t,\tau},R^{-1/2}\sum_{s=1}^{R-1}h'_{s+1})=0$. In addition, since A5 implies $\frac{(PR)^{1/2}}{T}=O(1)$, we immediately find that

$$\begin{aligned} \lim_{P,R\rightarrow\infty}Cov(P^{-1/2}\sum_{t=R}^{T-\tau}f_{t,\tau},P^{1/2}H(T)') &= \lim_{P,R\rightarrow\infty}\frac{P}{T}Cov(P^{-1/2}\sum_{t=R}^{T-\tau}f_{t,\tau},P^{-1/2}\sum_{s=R}^{T-1}h'_{s+1}) \\ &= \frac{\pi}{1+\pi}S'_{fh} \end{aligned}$$

and we obtain the desired result. Note that this does not depend on whether the recursive or rolling scheme is used.

(b-ii) We do the decomposition three distinct ways: once for the recursive, once for the rolling with $P<R$, and once for the rolling with $P\geq R$.

(b-ii-recursive) Let $a_{R,s} = \sum_{j=0}^{P-1}(R+s+j)^{-1}$and note that $P^{-1/2}\sum_{t=R}^{T-1} H(t)$ equals $(R/P)^{1/2}a_{R,0}(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}) + P^{-1/2}\sum_{s=R+1}^{T-1} a_{R,s}h_{s+1}$. And as above, $P^{1/2}H(T) = \frac{(PR)^{1/2}}{T}(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}) + \frac{P}{T}(P^{-1/2}\sum_{s=R}^{T-1} h_{s+1})$. Given A4, showing that $\lim_{P,R\to\infty} Cov(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}, P^{-1/2}\sum_{s=R}^{T-1} h_{s+1})$ and $\lim_{P,R\to\infty} Cov(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}, P^{-1/2}\sum_{s=R}^{T-1} a_{R,s}h_{s+1})$ are both zero is straightforward. In addition, since A5 implies $a_{R,0} \sim \ln(1+\pi)$ and $\frac{(PR)^{1/2}}{T}$ are $O(1)$ we find that

$$\begin{aligned}\lim_{P,R\to\infty} Cov(P^{-1/2}\sum_{t=R}^{T-\tau} H(t), P^{1/2}H(T)) &= \lim_{P,R\to\infty} \frac{R}{T}a_{R,0}Cov(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}, R^{-1/2}\sum_{s=1}^{R-1} h'_{s+1}) \\ &+ \lim_{P,R\to\infty} \frac{P}{T}Cov(P^{-1/2}\sum_{s=R}^{T-1} a_{R,s}h_{s+1}, P^{-1/2}\sum_{s=R}^{T-1} h'_{s+1}).\end{aligned}$$

Since $\lim_{P,R\to\infty} Cov(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}, R^{-1/2}\sum_{s=1}^{R-1} h_{s+1}) = \lim_{P,R\to\infty} Var(R^{-1/2}\sum_{s=1}^{R-1} h_{s+1})$ we immediately find that the first right-hand side term equals $(1+\pi)^{-1}\ln(1+\pi)$. That the second right-hand side term equals $\frac{\pi}{1+\pi}(1-\pi^{-1}\ln(1+\pi))$ is delineated in the proof of Lemma A6 in West (1996). Adding the two pieces together provides the desired result.

(b-ii-rolling, $P < R$) Define $\sum_{t=R}^{T-1} H(t) = R^{-1}\sum_{s=1}^{P-1} sh_s + \frac{P}{R}\sum_{s=P}^{R} h_s + R^{-1}\sum_{s=R+1}^{T-1}(P-s-R)h_s = A_1 + A_2 + A_3$ and $H(T) = T^{-1}\sum_{s=1}^{P-1} h_s + T^{-1}\sum_{s=P}^{R} h_s + T^{-1}\sum_{s=R+1}^{T-1} h_s = B_1 + B_2 + B_3$. Given A4, it is straightforward to show $\lim_{P,R\to\infty} Cov(A_i, B'_j) = 0$ all $i \neq j$ and hence $\lim_{P,R\to\infty} Cov(P^{-1/2}\sum_{t=R}^{T-\tau} H(t), P^{1/2}H(T)') = \sum_{j=1}^{3}\lim_{P,R\to\infty} Cov(A_j, B'_j)$. For the second term[13] it's clear that $\lim_{P,R\to\infty} Cov(A_2, B'_2) = \lim_{P,R\to\infty} \frac{P(R-P)}{RT}Cov((R-P)^{-1/2}\sum_{s=P}^{R} h_s, (R-P)^{-1/2}\sum_{s=P}^{R} h'_s) = \frac{\pi}{1+\pi}(1-\pi)S_{hh}$.

For the first and third terms a bit more detail is needed. For ease of presentation let h_s be a scalar and define $\gamma_j = Eh_sh_{s-j}$, $d_j = \sum_{i=1}^{P-1-j} i$, and $c_j = \sum_{i=1}^{P-1-j}(P-i)$. Direct multiplication and taking expectations implies $\lim_{P,R\to\infty} Cov(A_1, B_1) = (RT)^{-1}[\sum_{j=1}^{P-2}\gamma_j d_j + \sum_{j=1}^{P-2}\gamma_j c_j + \gamma_0 d_0]$. A5 and straightforward algebra imply $(RT)^{-1}d_0 = \frac{P^2}{RT}(P^{-2}\sum_{i=1}^{P-2} i) \to \frac{\pi^2}{2(1+\pi)}$. Since $\gamma_0 + 2\sum_{j=1}^{P-2}\gamma_j \to S_{hh}$ the result will follow if $P^{-2}|\sum_{j=1}^{P-2}\gamma_j(d_j - d_0)|$ and $P^{-2}|\sum_{j=1}^{P-2}\gamma_j(c_j - d_0)|$ are both $o(1)$. We show the former, the proof of the latter is very similar. To do so note that $d_j \geq \int_1^{P-1}(x-j)dx$ and $d_0 \leq \int_0^P xdx$. Integrating we obtain $|d_j - d_0| \leq |P + j(P - \frac{3}{2})|$ and hence $P^{-2}|\sum_{j=1}^{P-2}\gamma_j(d_j - d_0)| \leq P^{-2}\sum_{j=1}^{P-2}|\gamma_j||P + j(P-\frac{3}{2})| \leq P^{-1}\sum_{j=1}^{P-2}|\gamma_j| + P^{-1}\sum_{j=1}^{P-2}|\gamma_j||j|$. Since A4 implies $\sum_{j=1}^{P-2}|\gamma_j|$ and $\sum_{j=1}^{P-2}|\gamma_j||j|$ are $O(1)$ it is clear that $P^{-2}|\sum_{j=1}^{P-2}\gamma_j(d_j - d_0)| = o(1)$ and we

[13] Here we treat the case in which $|P-R|$ diverges as $P, R \to \infty$. When the difference is finite the second term is trivially $o_p(1)$. The distinction does not effect the ultimate formula for the asymptotic variance since, if $|P-R|$ diverges and $\pi = 1$, we still obtain the result that the second term is $o_p(1)$ (because the asymptotic variance of this term is zero).

conclude that $\lim_{P,R\to\infty} Cov(A_1, B_1) = \frac{\pi^2}{2(1+\pi)} S_{hh}$.

Moving to the third term, direct multiplication and taking expectations implies $\lim_{P,R\to\infty} Cov(A_3, B_3) = (RT)^{-1}[\sum_{j=1}^{P-2} \gamma_j d_j + \sum_{j=1}^{P-2} \gamma_j c_j + \gamma_0 d_0]$ if we redefine $c_j = \sum_{i=2}^{P-j} i$. Unsurprisingly, since this expansion is nearly identical to that for the first term, nearly identical arguments imply $\lim_{P,R\to\infty} Cov(A_3, B_3) = \frac{\pi^2}{2(1+\pi)} S_{hh}$.

Finally, if we add the three terms together we find that $\sum_{j=1}^{3} \lim_{P,R\to\infty} Cov(A_j, B_j') = \frac{\pi^2}{2(1+\pi)} S_{hh} + \frac{\pi}{1+\pi}(1-\pi) S_{hh} + \frac{\pi^2}{2(1+\pi)} S_{hh} = \frac{\pi}{1+\pi} S_{hh}$ and the proof is complete.

(b-ii-rolling, $P \geq R$) Define $\sum_{t=R}^{T-1} H(t) = R^{-1} \sum_{s=1}^{R} s h_s + \sum_{s=R+1}^{P} h_s + R^{-1} \sum_{s=1}^{R-1} s h_{T-s} = A_1 + A_2 + A_3$ and $H(T) = T^{-1} \sum_{s=1}^{R} h_s + T^{-1} \sum_{s=R+1}^{P} h_s + T^{-1} \sum_{s=1}^{R-1} h_{T-s} = B_1 + B_2 + B_3$. Given A4, it is straightforward to show $\lim_{P,R\to\infty} Cov(A_i, B_j') = 0$ all $i \neq j$ and hence $\lim_{P,R\to\infty} Cov(P^{-1/2} \sum_{t=R}^{T-\tau} H(t), P^{1/2} H(T)') = \sum_{j=1}^{3} \lim_{P,R\to\infty} Cov(A_j, B_j')$. For the second term[14] it's clear that $\lim_{P,R\to\infty} Cov(A_2, B_2') = \lim_{P,R\to\infty} \frac{(P-R)}{T} Cov((P-R)^{-1/2} \sum_{s=R+1}^{P} h_s, (P-R)^{-1/2} \sum_{s=R+1}^{P} h_s') = \frac{\pi-1}{1+\pi} S_{hh}$.

Once again, for the first and third terms a bit more detail is needed. For ease of presentation let h_s be a scalar and define $\gamma_j = E h_s h_{s-j}$, $d_j = \sum_{i=1}^{R-j} i$, and $c_j = \sum_{i=j+1}^{R} i$. Direct multiplication and taking expectations implies $\lim_{P,R\to\infty} Cov(A_1, B_1) = (RT)^{-1}[\sum_{j=1}^{R-1} \gamma_j d_j + \sum_{j=1}^{R-1} \gamma_j c_j + \gamma_0 d_0]$. A5 and straightforward algebra imply $(RT)^{-1} d_0 = \frac{R^2}{RT}(R^{-2} \sum_{i=1}^{R-1} i) \to \frac{1}{2(1+\pi)}$. Since $\gamma_0 + 2\sum_{j=1}^{R-1} \gamma_j \to S_{hh}$ the result will follow if $R^{-2} |\sum_{j=1}^{R-1} \gamma_j (d_j - d_0)|$ and $R^{-2} |\sum_{j=1}^{R-1} \gamma_j (c_j - d_0)|$ are both $o(1)$. We show the former, the proof of the latter is very similar. To do so note that $d_j \geq \int_1^{R-1} (x-j) dx$ and $d_0 \leq \int_0^R x dx$. Integrating we obtain $|d_j - d_0| \leq |R + j(R-2)|$ and hence $R^{-2} |\sum_{j=1}^{R-1} \gamma_j (d_j - d_0)| \leq R^{-2} \sum_{j=1}^{R-1} |\gamma_j| |R + j(R-2)| \leq R^{-1} \sum_{j=1}^{R-1} |\gamma_j| + R^{-1} \sum_{j=1}^{R-1} |\gamma_j| |j|$. Since A4 implies $\sum_{j=1}^{R-1} |\gamma_j|$ and $\sum_{j=1}^{R-1} |\gamma_j| |j|$ are $O(1)$ it is clear that $R^{-2} |\sum_{j=1}^{R-1} \gamma_j (d_j - d_0)| = o(1)$ and we conclude that $\lim_{P,R\to\infty} Cov(A_1, B_1) = \frac{1}{2(1+\pi)} S_{hh}$.

Moving to the third term, we again find that $\lim_{P,R\to\infty} Cov(A_3, B_3) = (RT)^{-1}[\sum_{j=1}^{R-2} \gamma_j d_j + \sum_{j=1}^{R-2} \gamma_j c_j + \gamma_0 d_0]$ if we redefine $d_j = \sum_{i=1}^{R-1-j} i$ and $c_j = \sum_{i=j+1}^{R-1} i$. Unsurprisingly, since this expansion is nearly identical to that for the first term, nearly identical arguments imply $\lim_{P,R\to\infty} Cov(A_3, B_3) = \frac{1}{2(1+\pi)} S_{hh}$.

Finally, if we add the three terms together we find that $\sum_{j=1}^{3} \lim_{P,R\to\infty} Cov(A_j, B_j') = \frac{1}{2(1+\pi)} S_{hh} + \frac{\pi-1}{1+\pi} S_{hh} + \frac{1}{2(1+\pi)} S_{hh} = \frac{\pi}{1+\pi} S_{hh}$ and the proof is complete.

[14] See the previous footnote.

References

Aastveit, Knut Are, Andrea Carriero, Todd E. Clark, and Massimiliano Marcellino (2014), "Have Standard VARs Remained Stable Since the Crisis?" Federal Reserve Bank of Cleveland Working Paper 14-11.

Baumeister, Christiane, and Lutz Kilian (2013), "Real-Time Analysis of Oil Price Risks Using Forecast Scenarios," *IMF Economic Review* 62, April 2014, 119-145.

Bernanke, Ben S., and Alan S. Blinder (1992), "The Federal Funds Rate and the Channels of Monetary Transmission," *American Economic Review* 82, 901-921.

Bruggemann, Ralf, Carsten Jentsch, and Carsten Trenkler (2014), "Inference in VARs with Conditional Heteroskedasticity of Unknown Form," manuscript, University of Mannheim.

Christiano, Lawrence J., Martin Eichenbaum, and Charles Evans (1996), "The Effects of Monetary Policy Shocks: Evidence from the Flow of Funds," *Review of Economics and Statistics* 78, 16-34.

Clarida, Richard H., and Coyle, Diane (1984), "Conditional Projection by Means of Kalman Filtering." Mimeo, National Bureau of Economic Research.

Clark, Todd E., and Taeyoung Doh (2014), "Evaluating Alternative Models of Trend Inflation," *International Journal of Forecasting* 30, 426-448.

Clark, Todd E., and Michael W. McCracken (2013), "Evaluating the Accuracy of Forecasts from Vector Autoregressions," in *Advances in Econometrics: VAR models in Macroeconomics — New Developments and Applications*, Thomas Fomby, Lutz Kilian, and Anthony Murphy (eds.), Emerald, 117-168.

Del Negro, Marco, and Frank Schorfheide (2013), "DSGE Model-Based Forecasting," in G. Elliott and A. Timmermann (eds.), *Handbook of Economic Forecasting*, vol. 2, Amsterdam: North-Holland.

Diebold, Francis X., and Roberto S. Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics* 13, 253-263.

Del Negro, Marco, Marc P. Giannoni, and Frank Schorfheide (2015), "Inflation in the Great Recession and New Keynesian Models," *American Economics Journal: Macroeconomics* 7, 168-196

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

Edge, Rochelle M., and Refet S. Gurkaynak (2010), "How Useful Are Estimated DSGE Model Forecasts for Central Bankers?" *Brookings Papers on Economic Activity* 41, 209-259.

Faust, Jon, and Jonathan H. Wright (2008), "Efficient Forecast Tests for Conditional Policy Forecasts," *Journal of Econometrics* 146, 293-303.

Faust, Jon, and Jonathan H. Wright (2013), "Forecasting Inflation," in *Handbook of Economic Forecasting*, volume 2 (North Holland).

Giannone, Domenico, Michele Lenza, and Giorgio Primiceri (2012), "Prior Selection for Vector Autoregressions," *Review of Economics and Statistics*, forthcoming.

Giannone, D., Lenza, M., Momferatou, D., and Onorante L. (2014), "Short-Term Inflation Projections: A Bayesian Vector Autoregressive Approach," *International Journal of Forecasting* 30, 635-644.

Goncalves, Silvia, and Lutz Kilian (2004), "Bootstrapping Autoregressions with Conditional Heteroskedasticity of Unknown Form," *Journal of Econometrics* 123, 89-120.

Gurkaynak, Refet S., Burcin Kisacikoglu, and Barbara Rossi (2013), "Do DSGE Models Forecast More Accurately Out-of-Sample than Reduced-Form Models?" in *Advances in Econometrics: VAR models in Macroeconomics — New Developments and Applications*, Thomas Fomby, Lutz Kilian, and Anthony Murphy (eds.), Emerald, 27-79.

Herbst, Edward, and Frank Schorfheide (2012), "Evaluating DSGE Model Forecasts of Comovements," *Journal of Econometrics* 171, 152-166.

Jarocinski, Marek (2010), "Conditional Forecasts and Uncertainty about Forecast Revisions in Vector Autoregressions," *Economics Letters* 108, 257-259.

Litterman, Robert B. (1986), "Forecasting with Bayesian Vector Autoregressions-Five Years of Experience," *Journal of Business and Economic Statistics* 4, 25-38.

Lutkepohl, Helmut (1991), *Introduction to Multiple Time Series*, Springer-Verlag Berlin, Heidelberg.

Meyer, Brent, and Saeed Zaman (2013), "It's Not Just for Inflation: The Usefulness of the Median CPI in BVAR Forecasting," FRB-Cleveland Working Paper 13-03.

Mincer, J. and Victor Zarnowitz (1969), "The Evaluation of Economic Forecasts," in J. Mincer (ed.), *Economic Forecasts and Expectations*, 3-46, National Bureau of Economic Research.

Newey, Whitney K., and Kenneth D. West (1987), "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica* 55, 703–708.

Schorfheide, Frank, and Dongho Song (2013), "Real Time Forecasting with a Mixed-Frequency VAR," NBER Working Paper 19712.

Smets, Frank, and Rafael Wouters (2007), "Shocks And Frictions In US Business Cycles: A Bayesian DSGE Approach, *American Economic Review* 97, 586-606.

Sims, Christopher A. (1982), "Policy Analysis with Econometric Models," *Brookings Papers on Economic Activity*, No. 1, 107-152.

Stine, Robert A. (1987), "Estimating Properties of Autoregressive Forecasts," *Journal of the American Statistical Association* 82, 1072-1078.

West, Kenneth D. (1996), "Asymptotic Inference About Predictive Ability," *Econometrica* 64, 1067-1084.

West, Kenneth D., and Michael W. McCracken (1998), "Regression-based Tests of Predictive Ability," *International Economic Review* 39, 817-840.

Table 1. Monte Carlo DGP coefficients

explanatory variable	$y_{1,t}$ equation	$y_{2,t}$ equation	$y_{3,t}$ equation
DGP 1 (size)			
$y_{1,t-1}$	0.50	0.00	
$y_{2,t-1}$	0.10	0.80	
intercept	0.00	0.00	
DGP 2 (size)			
$y_{1,t-1}$	0.234	0.029	0.059
$y_{1,t-2}$	0.164	-0.039	0.031
$y_{2,t-1}$	-0.134	0.575	0.038
$y_{2,t-2}$	-0.150	0.138	0.019
$y_{3,t-1}$	-0.057	0.200	1.006
$y_{3,t-2}$	-0.165	-0.184	-0.087
intercept	2.425	0.054	-0.110
DGP 3 (power)			
$y_{1,t-1}$, pre-break	0.50	0.00	
$y_{1,t-1}$, post-break	0.50	0.25	
$y_{2,t-1}$, pre-break	0.10	0.80	
$y_{2,t-1}$, post-break	0.10	0.40	
intercept, pre-break	0.00	0.00	
intercept, post-break	0.50	0.40	

Notes:
1. The table provides the coefficients of Monte Carlo DGPs 1-3. Other details of DGPs 1G and 4 are given in section 4.1.
2. The variance-covariance matrix of innovations and other aspects of the Monte Carlo design are described in section 4.1.

Table 2: Monte Carlo Results on Size, Minimum-MSE Conditioning, DGP 1
(nominal size = 10%)

test	*source of critical values*	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
1-step horizon					
bias, uncond.	Normal	0.115	0.132	0.118	0.123
bias, condit.	Normal	0.119	0.121	0.119	0.116
M-Z efficiency, uncond.	Normal	0.142	0.117	0.141	0.113
M-Z efficiency, condit.	Normal	0.170	0.132	0.146	0.114
F-W efficiency, condit.	Normal	0.147	0.102	0.126	0.101
equal MSE	Normal	0.008	0.005	0.057	0.026
bias, uncond.	bootstrap	0.118	0.122	0.128	0.114
bias, condit.	bootstrap	0.116	0.114	0.129	0.105
M-Z efficiency, uncond.	bootstrap	0.128	0.103	0.130	0.104
M-Z efficiency, condit.	bootstrap	0.121	0.095	0.132	0.103
F-W efficiency, condit.	bootstrap	0.129	0.104	0.127	0.100
equal MSE	bootstrap	0.107	0.100	0.135	0.098
2-step horizon					
bias, uncond.	Normal	0.116	0.126	0.130	0.121
bias, condit.	Normal	0.118	0.117	0.131	0.117
M-Z efficiency, uncond.	Normal	0.257	0.227	0.264	0.182
M-Z efficiency, condit.	Normal	0.205	0.174	0.196	0.148
F-W efficiency, condit.	Normal	0.240	0.188	0.235	0.162
equal MSE	Normal	0.022	0.006	0.086	0.043
bias, uncond.	bootstrap	0.119	0.114	0.120	0.114
bias, condit.	bootstrap	0.118	0.115	0.132	0.110
M-Z efficiency, uncond.	bootstrap	0.085	0.061	0.116	0.075
M-Z efficiency, condit.	bootstrap	0.114	0.089	0.117	0.090
F-W efficiency, condit.	bootstrap	0.085	0.064	0.110	0.074
equal MSE	bootstrap	0.074	0.046	0.115	0.079

Notes:
1. The data generating process is a bivariate VAR(1), with coefficients given in Table 1 and error variance matrix given in section 4.1.
2. For each artificial data set, forecasts of $y_{1,t+1}$ and $y_{1,t+2}$ are formed recursively using OLS estimates of a bivariate VAR(1). We consider both unconditional forecasts and conditional forecasts obtained under the minimum-MSE approach. The conditional forecasts of $y_{1,t+1}$ and $y_{1,t+2}$ are based on a condition of $\hat{y}^c_{2,t,\tau} = y_{2,t+\tau}$, $\tau = 1, 2$. These forecasts are then used to form bias, efficiency, and accuracy tests, detailed in sections 3 and 4.2.
3. R and P refer to the number of in–sample observations and forecasts, respectively.
4. In each Monte Carlo replication, the simulated test statistics are compared against Gaussian critical values and critical values obtained with the bootstrap of the VARs described in section 3.5.
5. The number of Monte Carlo simulations is 2000; the number of bootstrap draws is 499.

Table 3: Monte Carlo Results on Size, Minimum-MSE Conditioning, DGP 2
(nominal size = 10%)

		1-step horizon, variable 1				**1-step horizon, variable 2**			
test	*source of critical values*	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
bias, uncond.	Normal	0.145	0.129	0.130	0.132	0.139	0.133	0.136	0.134
bias, condit.	Normal	0.132	0.132	0.122	0.124	0.150	0.143	0.144	0.134
M-Z efficiency, uncond.	Normal	0.414	0.427	0.215	0.248	0.251	0.256	0.199	0.189
M-Z efficiency, condit.	Normal	0.482	0.499	0.232	0.284	0.320	0.344	0.228	0.230
F-W efficiency, condit.	Normal	0.403	0.407	0.234	0.240	0.234	0.251	0.196	0.193
equal MSE	Normal	0.035	0.018	0.066	0.038	0.143	0.155	0.111	0.103
bias, uncond.	bootstrap	0.120	0.104	0.127	0.101	0.109	0.091	0.123	0.103
bias, condit.	bootstrap	0.119	0.118	0.124	0.107	0.111	0.096	0.133	0.101
M-Z efficiency, uncond.	bootstrap	0.096	0.086	0.107	0.099	0.097	0.102	0.104	0.108
M-Z efficiency, condit.	bootstrap	0.085	0.066	0.100	0.089	0.094	0.105	0.100	0.103
F-W efficiency, condit.	bootstrap	0.087	0.080	0.099	0.096	0.097	0.102	0.093	0.110
equal MSE	bootstrap	0.079	0.048	0.110	0.090	0.054	0.043	0.086	0.069
		2-step horizon, variable 1				**2-step horizon, variable 2**			
bias, uncond.	Normal	0.141	0.131	0.143	0.125	0.142	0.133	0.150	0.135
bias, condit.	Normal	0.124	0.127	0.141	0.131	0.142	0.133	0.154	0.128
M-Z efficiency, uncond.	Normal	0.312	0.292	0.211	0.206	0.315	0.310	0.294	0.242
M-Z efficiency, condit.	Normal	0.399	0.393	0.245	0.251	0.394	0.388	0.310	0.290
F-W efficiency, condit.	Normal	0.305	0.288	0.251	0.214	0.290	0.289	0.296	0.241
equal MSE	Normal	0.060	0.037	0.093	0.059	0.073	0.066	0.093	0.079
bias, uncond.	bootstrap	0.116	0.107	0.126	0.104	0.107	0.091	0.123	0.105
bias, condit.	bootstrap	0.115	0.116	0.122	0.113	0.108	0.097	0.126	0.098
M-Z efficiency, uncond.	bootstrap	0.090	0.089	0.104	0.097	0.088	0.098	0.102	0.100
M-Z efficiency, condit.	bootstrap	0.081	0.077	0.102	0.091	0.091	0.092	0.101	0.102
F-W efficiency, condit.	bootstrap	0.082	0.084	0.103	0.090	0.087	0.097	0.102	0.110
equal MSE	bootstrap	0.070	0.043	0.099	0.084	0.058	0.048	0.100	0.075
		4-step horizon, variable 1				**4-step horizon, variable 2**			
bias, uncond.	Normal	0.138	0.129	0.169	0.142	0.141	0.126	0.171	0.136
bias, condit.	Normal	0.132	0.131	0.168	0.144	0.147	0.118	0.177	0.136
M-Z efficiency, uncond.	Normal	0.313	0.274	0.285	0.217	0.561	0.554	0.462	0.418
M-Z efficiency, condit.	Normal	0.358	0.345	0.285	0.254	0.499	0.493	0.392	0.361
F-W efficiency, condit.	Normal	0.311	0.264	0.300	0.225	0.530	0.524	0.464	0.415
equal MSE	Normal	0.097	0.088	0.154	0.105	0.056	0.042	0.116	0.062
bias, uncond.	bootstrap	0.104	0.110	0.120	0.114	0.106	0.093	0.112	0.100
bias, condit.	bootstrap	0.103	0.107	0.120	0.111	0.110	0.090	0.117	0.102
M-Z efficiency, uncond.	bootstrap	0.066	0.059	0.099	0.064	0.061	0.061	0.092	0.069
M-Z efficiency, condit.	bootstrap	0.077	0.080	0.105	0.086	0.072	0.081	0.094	0.082
F-W efficiency, condit.	bootstrap	0.057	0.061	0.100	0.062	0.065	0.062	0.081	0.084
equal MSE	bootstrap	0.048	0.032	0.095	0.073	0.044	0.036	0.090	0.068

Notes:
1. The data generating process is a trivariate VAR(2), with coefficients given in Table 1 and error variance matrix given in section 4.1.
2. For each artificial data set, forecasts of $y_{1,t+\tau}$ and $y_{2,t+\tau}$ are formed recursively using OLS estimates of a trivariate VAR(2) and an iterative approach to computing multi-step forecasts. The conditional forecasts of $y_{1,t+\tau}$ and $y_{2,t+\tau}$ (for τ = 1, 2, and 4) are based on a condition of $\hat{y}^c_{3,t,\tau} = y_{3,t+\tau}$, $\tau = 1, \ldots, 4$.
3. See the notes to Table 2.

Table 4: Monte Carlo Results on Size, Minimum-MSE Conditioning, DGP 1, rolling estimation

(nominal size = 10%)

test	*source of critical values*	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
1-step horizon					
bias, uncond.	Normal	0.018	0.009	0.099	0.057
bias, condit.	Normal	0.022	0.013	0.100	0.065
M-Z efficiency, uncond.	Normal	0.192	0.233	0.142	0.087
M-Z efficiency, condit.	Normal	0.212	0.269	0.142	0.088
F-W efficiency, condit.	Normal	0.160	0.148	0.122	0.073
equal MSE	Normal	0.026	0.013	0.059	0.037
bias, uncond.	bootstrap	0.116	0.097	0.123	0.114
bias, condit.	bootstrap	0.117	0.095	0.138	0.110
M-Z efficiency, uncond.	bootstrap	0.132	0.098	0.141	0.112
M-Z efficiency, condit.	bootstrap	0.122	0.076	0.143	0.106
F-W efficiency, condit.	bootstrap	0.128	0.076	0.132	0.103
equal MSE	bootstrap	0.104	0.091	0.134	0.103
2-step horizon					
bias, uncond.	Normal	0.013	0.004	0.112	0.058
bias, condit.	Normal	0.018	0.009	0.115	0.064
M-Z efficiency, uncond.	Normal	0.567	0.711	0.276	0.236
M-Z efficiency, condit.	Normal	0.324	0.417	0.200	0.136
F-W efficiency, condit.	Normal	0.435	0.508	0.245	0.195
equal MSE	Normal	0.040	0.017	0.090	0.050
bias, uncond.	bootstrap	0.118	0.101	0.126	0.117
bias, condit.	bootstrap	0.116	0.088	0.132	0.112
M-Z efficiency, uncond.	bootstrap	0.079	0.046	0.115	0.066
M-Z efficiency, condit.	bootstrap	0.099	0.059	0.126	0.095
F-W efficiency, condit.	bootstrap	0.079	0.038	0.099	0.077
equal MSE	bootstrap	0.088	0.061	0.113	0.088

Notes:
1. In these experiments, forecasts of $y_{1,t+1}$ and $y_{1,t+2}$ are formed recursively using OLS estimates of a bivariate VAR(1).
2. See the notes to Table 2.

Table 5: Monte Carlo Results on Size, Minimum-MSE Conditioning, DGP 1G
(nominal size = 10%)

test	*source of critical values*	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
1-step horizon					
bias, uncond.	Normal	0.118	0.120	0.114	0.114
bias, condit.	Normal	0.117	0.112	0.118	0.113
M-Z efficiency, uncond.	Normal	0.147	0.141	0.140	0.117
M-Z efficiency, condit.	Normal	0.192	0.173	0.155	0.147
F-W efficiency, condit.	Normal	0.134	0.130	0.134	0.122
equal MSE	Normal	0.024	0.013	0.082	0.045
bias, uncond.	bootstrap	0.109	0.124	0.129	0.108
bias, condit.	bootstrap	0.120	0.114	0.127	0.101
M-Z efficiency, uncond.	bootstrap	0.110	0.114	0.135	0.103
M-Z efficiency, condit.	bootstrap	0.117	0.102	0.120	0.097
F-W efficiency, condit.	bootstrap	0.114	0.102	0.128	0.108
equal MSE	bootstrap	0.114	0.099	0.122	0.114
2-step horizon					
bias, uncond.	Normal	0.118	0.111	0.126	0.116
bias, condit.	Normal	0.113	0.105	0.130	0.110
M-Z efficiency, uncond.	Normal	0.267	0.246	0.237	0.205
M-Z efficiency, condit.	Normal	0.227	0.199	0.218	0.180
F-W efficiency, condit.	Normal	0.224	0.223	0.217	0.183
equal MSE	Normal	0.028	0.021	0.095	0.061
bias, uncond.	bootstrap	0.110	0.123	0.130	0.108
bias, condit.	bootstrap	0.115	0.115	0.124	0.107
M-Z efficiency, uncond.	bootstrap	0.064	0.062	0.107	0.073
M-Z efficiency, condit.	bootstrap	0.107	0.092	0.123	0.095
F-W efficiency, condit.	bootstrap	0.074	0.068	0.106	0.076
equal MSE	bootstrap	0.079	0.057	0.106	0.083

Notes:
1. The data generating process is a bivariate VAR(1) with GARCH, with coefficients given in Table 1 and and section 4.1.
2. See the notes to Table 2.

Table 6: Monte Carlo Results on Power, Minimum-MSE Conditioning, DGP 3
(nominal size = 10%)

test	*source of critical values*	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
1-step horizon					
bias, uncond.	bootstrap	0.628	0.642	0.740	0.850
bias, condit.	bootstrap	0.444	0.481	0.587	0.694
M-Z efficiency, uncond.	bootstrap	0.244	0.315	0.113	0.158
M-Z efficiency, condit.	bootstrap	0.134	0.126	0.120	0.100
F-W efficiency, condit.	bootstrap	0.136	0.159	0.089	0.085
equal MSE	bootstrap	0.115	0.104	0.132	0.119
2-step horizon					
bias, uncond.	bootstrap	0.632	0.657	0.738	0.850
bias, condit.	bootstrap	0.510	0.535	0.643	0.748
M-Z efficiency, uncond.	bootstrap	0.230	0.314	0.108	0.173
M-Z efficiency, condit.	bootstrap	0.104	0.113	0.099	0.079
F-W efficiency, condit.	bootstrap	0.115	0.127	0.079	0.071
equal MSE	bootstrap	0.104	0.064	0.171	0.149

Notes:
1. The data generating process is a bivariate VAR(1), with coefficient breaks, using the coefficient values given in Table 1 and the error variance matrix given in section 4.1. In each experiment, the coefficient break occurs in period R+1.
2. See the notes to Table 2.

Table 7: Monte Carlo Results on Power, Minimum-MSE Conditioning, DGP 4
(nominal size = 10%)

test	*source of critical values*	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
1-step horizon					
bias, uncond.	bootstrap	0.141	0.120	0.122	0.106
bias, condit.	bootstrap	0.096	0.103	0.097	0.069
M-Z efficiency, uncond.	bootstrap	0.129	0.104	0.128	0.114
M-Z efficiency, condit.	bootstrap	0.199	0.119	0.609	0.466
F-W efficiency, condit.	bootstrap	0.147	0.103	0.169	0.178
equal MSE	bootstrap	0.490	0.398	0.945	0.876
2-step horizon					
bias, uncond.	bootstrap	0.153	0.140	0.128	0.115
bias, condit.	bootstrap	0.091	0.103	0.099	0.072
M-Z efficiency, uncond.	bootstrap	0.082	0.067	0.127	0.086
M-Z efficiency, condit.	bootstrap	0.335	0.198	0.757	0.723
F-W efficiency, condit.	bootstrap	0.112	0.081	0.147	0.130
equal MSE	bootstrap	0.459	0.418	0.854	0.822

Notes:
1. The data generating process is a bivariate VAR(1), with a break in the error correlation described in section 4.1. In each experiment, the coefficient break occurs in period R+1.
2. See the notes to Table 2.

Table 8. Tests of unconditional and conditional (min.-MSE) forecasts from 7-variable VAR, 1991-2007

(significant tests bolded; 10% bootstrap critical values in parentheses)

GDP growth	τ=1	τ=2	τ=4	τ=8
bias, uncond.	**-3.603**	**-3.424**	**-3.019**	**-2.258**
	(-1.687, 1.973)	(-2.002, 1.942)	(-1.878, 2.002)	(-1.806, 2.102)
bias, condit.	**-2.970**	**-2.265**	**-2.756**	-1.983
	(-1.616, 1.891)	(-1.636, 1.679)	(-1.749, 1.830)	(-2.235, 2.238)
M-Z efficiency, uncond.	-3.429	**-4.866**	-2.574	-1.844
	(-4.619, -0.204)	(-4.732, -0.040)	(-6.021, 0.155)	(-5.845, 0.928)
M-Z efficiency, condit.	-3.953	**-5.033**	-2.025	-4.964
	(-4.877, -0.243)	(-4.821, -0.144)	(-4.774, 0.392)	(-5.728, 1.298)
F-W efficiency, condit.	-3.571	**-6.950**	-2.590	-1.991
	(-4.660, -0.161)	(-4.857, -0.088)	(-5.004, 0.301)	(-4.863, 1.521)
Investment growth	τ=1	τ=2	τ=4	τ=8
bias, uncond.	-1.062	-1.368	-1.013	-0.613
	(-1.675, 1.737)	(-1.820, 1.936)	(-1.686, 1.981)	(-1.795, 2.064)
bias, condit.	-0.717	-0.613	-0.619	-0.840
	(-1.647, 1.706)	(-1.646, 1.708)	(-1.504, 1.715)	(-1.994, 2.230)
M-Z efficiency, uncond.	-3.290	-3.599	-4.604	-1.123
	(-4.758, 0.000)	(-4.642, 0.232)	(-4.711, -0.014)	(-6.513, 1.348)
M-Z efficiency, condit.	-2.962	**-3.676**	-2.434	-1.861
	(-4.155, -0.083)	(-3.618, -0.027)	(-3.655, 0.290)	(-5.580, 1.529)
F-W efficiency, condit.	-3.229	**-5.666**	**-5.955**	-0.790
	(-4.106, 0.040)	(-4.121, 0.168)	(-4.386, 0.172)	(-5.230, 1.884)
Inflation	τ=1	τ=2	τ=4	τ=8
bias, uncond.	-1.656	-1.968	-2.000	-3.459
	(-2.581, 1.531)	(-2.768, 1.494)	(-3.117, 1.581)	(-4.171, 2.397)
bias, condit.	-1.340	-1.622	-1.196	-1.290
	(-2.685, 1.750)	(-2.873, 1.821)	(-3.189, 1.954)	(-3.952, 3.228)
M-Z efficiency, uncond.	**-5.996**	**-7.000**	-4.901	-5.852
	(-4.794, 0.226)	(-5.632, 0.275)	(-5.090, 0.367)	(-6.933, 0.908)
M-Z efficiency, condit.	**-5.395**	**-6.164**	**-5.523**	**-8.453**
	(-4.767, 0.609)	(-5.074, 0.683)	(-4.949, 0.554)	(-6.828, 0.854)
F-W efficiency, condit.	**-5.288**	**-7.291**	**-5.477**	-6.065
	(-4.761, 0.460)	(-5.435, 0.614)	(-5.376, 0.567)	(-6.719, 1.209)

Notes:

1. As described in section 5.1, forecasts of growth in GDP, growth in private fixed investment, and inflation in the GDP price index (all defined at annualized rates) are obtained from recursive OLS estimates of a 7-variable VAR. The forecasts included are unconditional and minimum-MSE conditional. At each forecast origin t, the conditions imposed are that, over an eight quarter forecast horizon from $t+1$ through $t+8$, the federal funds rate take its actual values over the period.

2. The bias, efficiency, and MSE accuracy tests detailed in section 4.2 are compared against standard normal critical values and critical values obtained with the bootstrap described in section 3.5. The number of bootstrap draws is 999.

Table 9. Accuracy of unconditional and conditional (min.-MSE) forecasts from 7-variable VAR, 1991-2007

(significant tests bolded; 10% bootstrap critical values in parentheses)

GDP growth	τ=1	τ=2	τ=4	τ=8
unconditional MSE (RMSE)	4.495 (2.120)	5.273 (2.296)	5.504 (2.346)	4.533 (2.129)
conditional MSE (RMSE)	4.351 (2.086)	5.028 (2.242)	4.050 (2.012)	5.203 (2.281)
$\text{MSE}_U - \text{MSE}_C$	0.144	0.245	1.454	-0.670
population estimate of $\text{MSE}_U - \text{MSE}_C$	0.710	1.293	2.381	2.053
MSE t-test	**-2.391**	-2.337	-1.490	-3.740
10% bootstrap crit. vals.	(-2.230, 0.814)	(-2.380, 1.127)	(-3.119, 1.633)	(-6.494, 1.247)
Investment growth	τ=1	τ=2	τ=4	τ=8
unconditional MSE (RMSE)	35.231 (5.936)	36.387 (6.032)	46.746 (6.837)	32.976 (5.742)
conditional MSE (RMSE)	31.229 (5.588)	33.255 (5.767)	34.962 (5.913)	36.411 (6.034)
$\text{MSE}_U - \text{MSE}_C$	4.002	3.132	11.784	-3.435
population estimate of $\text{MSE}_U - \text{MSE}_C$	7.973	13.109	16.686	12.616
MSE t-test	-1.841	-3.104	-1.447	-3.229
10% bootstrap crit. vals.	(-2.324, 1.372)	(-3.122, 1.555)	(-3.681, 1.129)	(-4.524, 1.824)
Inflation	τ=1	τ=2	τ=4	τ=8
unconditional MSE (RMSE)	0.733 (0.856)	0.941 (0.970)	1.065 (1.032)	2.000 (1.414)
conditional MSE (RMSE)	0.645 (0.803)	0.698 (0.836)	0.904 (0.951)	1.835 (1.355)
$\text{MSE}_U - \text{MSE}_C$	0.088	0.243	0.161	0.165
population estimate of $\text{MSE}_U - \text{MSE}_C$	0.062	0.193	0.402	1.051
MSE t-test	0.652	0.554	-2.146	**-4.784**
10% bootstrap crit. vals.	(-1.600, 1.109)	(-1.485, 1.175)	(-2.415, 1.288)	(-3.681, 1.741)

Notes:
1. The table reports forecast MSEs (and RMSEs) for the 1991-2007 sample. It also includes estimates of the population difference in MSEs (unconditional less conditional) based on the full sample VAR estimates, computed as described in section 3.4.
2. See the notes to Table 8.