Evaluating Conditional Forecasts from Vector Autoregressions

by

Todd E. Clark and Michael W. McCracken

NBER/NSF Time Series Conference
September 26, 2015

The Motivation

- Central bank forecasts are often conditional on future monetary policy
  - GB/TB: hold the funds rate path constant over the next two years
  - FOMC: "appropriate (future) monetary policy"

- Bank Stress-testing procedures include conditional forecasting
  - Effect of severely adverse macroeconomic scenario on net interest margins, net charge-offs, etc.

<u>Punchline</u>: Conditional forecasts are common and important

The Problem

- The forecast evaluation literature concentrates on <u>un</u>conditional forecasts
  - <u>Unconditional</u>: Only uses known time t information.
  - <u>Conditional</u>: Also includes future or hypothesized-future information.

- Exceptions that evaluate conditional forecasts
  - <u>Modesty</u>: Doan, Littermann, & Sims (1984), Jorda & Marcellino (2010)
  - <u>Efficiency</u>: Faust & Wright (2008)
  - <u>Accuracy</u>: Herbst & Schoerfheide (2012)

<u>Punchline</u>: The forecast evaluation literature is rich for unconditional forecasts, but underdeveloped for conditional forecasts

Why the Lack of Methods for Evaluating Conditional Forecasts?

- Conditional forecasts depend on two pieces of information
  - the model
  - the conditioning path/scenario

- Disentangling the two effects is not obvious.
  - Is my model bad at producing conditional forecasts?
  - Is the scenario perverse/logically inconsistent?

- Perhaps worse, a badly misspecified model and perverse scenario can interact and give forecasts that look "ok."

<u>Punchline</u>: In general it seems impossible to separate perverse conditioning from a bad model.

A simpler, informative approach.

- Rather than try to disentangle the two effects, fix one and evaluate the other.
  - We focus on the models *ability* to construct good conditional forecasts
  - We do not evaluate scenarios per se

- Perform a sequence of pseudo-out-of-sample conditional forecasts of "y"
  - At each forecast origin $t = R, \ldots, R + P - \tau = T$, *condition on the future realized values* of those variables in the scenario (let's call them x).
  - If the model is "good" at constructing conditional forecasts, these conditional forecast errors should exhibit certain properties.

<u>Punchline</u>: Using pseudo out-of-sample methods, we develop tools for understanding when conditional forecasts from VARs should be "good."

What We Do

- Establish asymptotic normality of tests of bias, efficiency, and "equal accuracy" for conditional forecasts from VARs.

- Theory related to West (1996) and West & McCracken (1998)

- Discuss inference: N(0,1) cv's vs. Bootstrap cv's.

- Provide Monte Carlo evidence on Size and Power of the tests

- Investigate the 7-variable VAR from Smets and Wouters (2007) when conditioned on future values of the Fed Funds Rate. (omitted for time)

A Simple Example

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

- Want to forecast y, τ = 1- and 2-steps ahead given $\hat{x}^c_{t,1}$ and $\hat{x}^c_{t,2}$.

- Define $\hat{y}^c_{t,\tau}$ and $\hat{y}^u_{t,\tau}$ as the τ-step <u>c</u>onditional/<u>u</u>nconditional forecasts of y.

Simple Example (continued)

- Reduced form: $\hat{y}^c_{t,1} = \hat{y}^u_{t,1} + \hat{\rho}_t(\hat{x}^c_{t,1} - \hat{x}^u_{t,1})$

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

- Policy shock: $\hat{y}^c_{t,1} = \hat{y}^u_{t,1}$

$$\hat{y}^c_{t,2} = \hat{y}^u_{t,2} + \hat{b}_t(\hat{x}^c_{t,1} - \hat{x}^u_{t,1})$$

- In each case the τ-step conditional forecasts take the form

$$\hat{y}^c_{t,\tau} = \hat{y}^u_{t,\tau} + \sum_{j=1}^{m} \hat{\gamma}_{t,j}(\hat{x}^c_{t,j} - \hat{x}^u_{t,j})$$

some finite conditioning horizon $m > 0$ and constants γ_j.

Now Condition on Future Values of x

- If we let $\hat{x}^c_{t,j} = x_{t+j}$ we obtain

$$\hat{y}^c_{t,\tau} = \hat{y}^u_{t,\tau} + \sum_{j=1}^{m} \hat{\gamma}_{t,j}(x_{t+j} - \hat{x}^u_{t,j})$$

and hence

$$\hat{u}^c_{t,\tau} = \hat{u}^u_{t,\tau} - \sum_{j=1}^{m} \hat{\gamma}_{t,j}\hat{v}^u_{t,j}$$

<u>Punchline</u>: The problem of evaluating conditional forecasts is now a problem of evaluating a linear function of unconditional forecasts.

- Key distinction:
  - $\hat{u}^u_{t,\tau}$ depends on lagged dynamics of model
  - $\hat{u}^c_{t,\tau}$ depends on lagged <u>&</u> contemporaneous dynamics of model

Properties of Good Forecast Errors: Known Future $x's$

$$\hat{u}^c_{t,\tau} = \hat{u}^u_{t,\tau} - \sum_{j=1}^{m} \hat{\gamma}_{t,j} \hat{v}^u_{t,j}$$

- <u>Zero bias</u>: $Eu^c_{t,\tau} = 0$

- <u>Mincer-Zarnowitz efficiency</u>: $Eu^c_{t,\tau} \hat{y}^c_{t,\tau} = 0$
  - Holds for reduced-form, but not policy-shock conditioning
  - Faust-Wright (2008) solve the problem for policy-shock conditioning

- <u>Equal accuracy</u>: For some $k > 0$, $E(u^c_{t,\tau})^2 - E(u^u_{t,\tau})^2 + k = 0$ under reduced-form conditioning.

Test Statistics: $H_0 : \beta = 0$

- <u>Zero bias</u>: $\hat{u}^c_{t,\tau} = \beta + error$

- <u>Mincer-Zarnowitz efficiency</u>: $\hat{u}^c_{t,\tau} = \alpha + \beta \hat{y}^c_{t,\tau} + error$
  - Holds for reduced-form, but not policy-shock conditioning

- <u>Faust-Wright efficiency</u>: $\hat{u}^c_{t,\tau} = \alpha + \gamma(\hat{y}^u_{t,\tau} - \hat{y}^c_{t,\tau}) + \beta \hat{y}^c_{t,\tau} + error$
  - Holds for both reduced-form and policy-shock conditioning

An Equal Accuracy Test

- Good reduced-form conditional forecasts should have (weakly) smaller MSE than the corresponding unconditional forecasts.

- Let $\Psi_j \Sigma^{1/2}$ denote the matrix of orthogonalized IRFs after j periods. Define

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

and let $\tilde{D}$ denote the matrix formed by stacking the rows of D associated with a conditioning restriction

- $E(u_{1,t,\tau}^{u})^2 - E(u_{1,t,\tau}^{c})^2 = k(\phi) = \iota_1' D\tilde{D}'(\tilde{D}\tilde{D}')^{-1}\tilde{D}D'\iota_1$

A Statistic for Equal Accuracy

- $\beta = E(u_{1,t,\tau}^{c})^2 - E(u_{1,t,\tau}^{u})^2 + k(\phi)$

- $\hat{\beta} = P^{-1}\sum_{t=R}^{T-\tau}(\hat{u}_{1,t,\tau}^{c})^2 - (\hat{u}_{1,t,\tau}^{u})^2 + k(\hat{\phi}_T)$

- Important to note that the forecast errors depend on a sequence of re-estimated parameters $\hat{\phi}_t$ while the centering constant is constructed using the full sample. Asymptotics do not follow directly from West (1996).

Asymptotics

- Assumptions:
  - $(n \times 1)$ finite order VAR estimated by OLS yielding $\hat{\Lambda}_t$
  - residuals ~ mds and are used to estimate residual variance matrix $\hat{\Sigma}_t$
  - observables are fourth-order stationary, mixing, have 8+ moments
  - $\hat{y}^c_{1,t,\tau} = \hat{y}^u_{1,t,\tau} + \sum_{i=1}^{n} \sum_{j=1}^{m} \hat{\gamma}_{i,t,j} (\hat{x}^c_{i,t,j} - \hat{x}^u_{i,t,j})$
  - $\hat{\gamma}_{1,t,\tau} = 0$

- An application of the theory in West (1996)
  - Let $\hat{\phi}_t = (vec(\hat{\Lambda}_t)', vech(\hat{\Sigma}_t)')'$
  - $T^{1/2}(\hat{\phi}_T - \phi) = BH(T) + o_{a.s}(1)$
  - $P^{1/2}\hat{\beta} = P^{-1/2} \sum_{t=R}^{T} f(Z_t, \hat{\phi}_t) + o_p(1)$ or
    $P^{1/2}\hat{\beta} = P^{-1/2} \sum_{t=R}^{T} f(Z_t, \hat{\phi}_t) - k(\hat{\phi}_T) + o_p(1).$

Theorems

- Let $F = E(\partial f(Z_t, \phi) / \partial \phi)$, $K = \partial k(\phi) / \partial \phi$, $\lim_{P,R \to \infty} P / R = \pi$.

Theorems 1 and 2: For each of the zero bias, Mincer-Zarnowitz, Faust-Wright regressions $P^{1/2} \hat{\beta} \to N(0, \Omega)$ with

$$\Omega = S_{ff} + 2\lambda_{fh} FBS'_{fh} + \lambda_{hh} FBS_{hh} B'F'$$

Theorem 3: For the Equal Accuracy test, $P^{1/2} \hat{\beta} \to N(0, \Omega)$ with

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

Inference

- Estimate Ω directly and use standard normal critical values for $P^{1/2}\hat{\beta} / \hat{\Omega}^{1/2}$
  - Ω simplifies immensely only if $P / R \sim 0$
  - For bias and simple efficiency, $F_{\Sigma} = 0$
  - For MZ/FW efficiency and equal MSE, the F term is a mess

- Instead consider bootstrap of $P^{1/2}\hat{\beta} / \hat{S}_{ff}^{1/2}$.
  - Standard wild recursive VAR bootstrap doesn't work because it doesn't get the asymptotic distribution of the residual variance right.
  - iid recursive VAR bootstrap works but is restrictive.
  - We use the moving block, residual-based, recursive-VAR bootstrap due to Bruggemann, Carlsten, and Trenkler (2014).

Description of Monte Carlo

- $\begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} 0.5 & 0.10 \\ 0 & 0.80 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ x_{t-1} \end{pmatrix} + \begin{pmatrix} u_t \\ v_t \end{pmatrix}, \Sigma = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$

- $\tau = 1, 2, 4$, Conditional on known x_{t+1} and x_{t+2}

- Recursive and rolling forecasts

- N(0,1) vs. bootstrap-based critical values.

- A few in-sample (R = 50, 100) and out-of-sample (P = 100, 150) sizes

- Bias, MZ-efficiency, and equal accuracy.

- Size and Power
  - Power from unmodeled breaks in Λ or Σ.

Size of Tests

(nominal size = 10%, forecast & conditioning horizon = 2)

	CVs	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
Recursive					
bias, uncond.	BS	0.119	0.114	0.120	0.114
bias, condit.	BS	0.118	0.115	0.132	0.110
M-Z efficiency, unc.	BS	0.085	0.061	0.116	0.075
M-Z efficiency, con.	BS	0.114	0.089	0.117	0.090
equal MSE	BS	0.074	0.046	0.115	0.079
Rolling					
bias, uncond.	BS	0.118	0.101	0.126	0.117
bias, condit.	BS	0.116	0.088	0.132	0.112
M-Z efficiency, unc.	BS	0.079	0.046	0.115	0.066
M-Z efficiency, con.	BS	0.099	0.059	0.126	0.095
equal MSE	BS	0.088	0.061	0.113	0.088

Power of Tests, Recursive

(nominal size = 10%, forecast & conditioning horizon = 2)

Break in Λ	CVs	R=50 P=100	R=50 P=150	R=100 P=50	R=100 P=100
bias, unc.	BS	0.632	0.657	0.738	0.850
bias, con.	BS	0.510	0.535	0.643	0.748
M-Z efficiency, unc.	BS	0.230	0.314	0.108	0.173
M-Z efficiency, con.	BS	0.104	0.113	0.099	0.079
equal MSE	BS	0.104	0.064	0.171	0.149
Break in Σ					
bias, unc.	BS	0.153	0.140	0.128	0.115
bias, con.	BS	0.091	0.103	0.099	0.072
M-Z efficiency, unc.	BS	0.082	0.067	0.127	0.086
M-Z efficiency, con.	BS	0.335	0.198	0.757	0.723
equal MSE	BS	0.459	0.418	0.854	0.822

Conclusion

- Conditional Forecasts are common in central banking
  - GB, FOMC forecasts
  - Stress testing forecasts are conditional on severely adverse scenario

- And yet there are very few tools for assessing their quality

- In this paper we are interested in developing tools for understanding when conditional forecasts from VARs should be "good."

- We provide analytical, simulation, and empirical evidence on the quality of VARs used to construct conditional forecasts.