Modeling Time-Varying Uncertainty of Multiple-Horizon Forecast Errors

Todd E. Clark [1] **Michael W. McCracken** [2] **Elmar Mertens** [3]

[1]Federal Reserve Bank of Cleveland

[2]Federal Reserve Bank of St. Louis

[3]Bank for International Settlements

September 2017

Introduction

High level question:

Given history of judgmental point forecasts $E_t y_{t+h}$, for multiple horizons h, how can we estimate uncertainty $\text{Var}_t\, y_{t+h}$ that may be time-varying?

Introduction

In central bank communications on monetary policy, forecasts and forecast uncertainty play prominent roles

- Forecasts are typically judgmental and not entirely model-based
- Forecast fan charts in monetary policy reports

Central banks commonly use historical forecast errors to measure forecast uncertainty

- Examples: Reserve Bank of Australia, European Central Bank, Federal Reserve

Introduction

Example of Federal Reserve forecasts, in the Summary of Economic Projections (SEP)

- Point forecasts for real activity, inflation and interest rates
- Horizon: current year and up to two future calendar years

Treatment of uncertainty

- Qualitative assessments
- Table of historical RMSEs [Reifschneider & Tulip (2007, 2017)]
  - Based on historical forecast errors from variety of sources
  - Use 20-year MSE as estimate of forecast error variance
  - Regularly updated
- Since March 2017: fan charts using those RMSEs

Introduction

Historical uncertainty is commonly treated as constant

- May use a rolling window to accommodate some change over time: Federal Reserve's SEP
- Some central banks use a judiciously chosen sample period

Yet VAR and DSGE studies suggest significant time variation in forecast error variances: stochastic volatility

- Cogley & Sargent (2005), Primiceri (2005), D'Agostino, Gambetti & Giannone (2013), Clark & Ravazzolo (2015), Carriero, Clark & Marcellino (2016)
- Justiniano & Primiceri (2008), Diebold, Schorfheide & Shin (2016)
- SV improves density forecasts

Introduction

General challenge to SV with forecasts from a central bank or from a survey (e.g., SPF):

- Available forecasts and errors span multiple horizons, with overlap
- No such SV model exists; typical time series model is specified at a one-step ahead horizon, with multi-step errors inferred from the recursive nature of the parametric model

Introduction

We develop a multiple-horizon specification of SV for forecast errors from sources such as SPF

- Key to solution: decomposition of multi-step forecast error into sums of forecast updates
- Our approach yields confidence bands around forecasts that allow for variation over time in the width of the confidence bands
- Explicit modeling of time variation of volatility eliminates the need for somewhat arbitrary judgments of sample stability

We estimate the model with standard Bayesian methods for multivariate SV specifications

- Gibbs sampler (Primiceri 2005)
- Posterior forecast density

Introduction

Results from SPF data:

- We estimate the model with the full history of data to document considerable historical variation in forecast error variances
  - GDP growth, unemployment, inflation, and short-term interest rate
- We produce pseudo-real time estimates of forecast uncertainty and evaluate density forecasts implied by the SPF errors and our estimated uncertainty bands
  - Interval forecasts and CRPS
- Our proposed approach yields uncertainty estimates more accurate than those obtained using simple historical RMSEs

Results qualitatively similar with Greenbook forecasts

Introduction

Related literature: survey forecasts

- Liu & Lahiri (2006), Lahiri & Sheng (2010)
- D'Amico & Orphanides (2008), Clements (2014/16), Boero, Smith & Wallis (2015)
- Ball & Croushore (2003), Rudebusch & Williams (2009)
- Coibion & Gorodnichenko (2012/15), Mertens & Nason (2015)

Related literature: uncertainty based on past forecast errors

- Reifschneider & Tulip (2007, 2017), Knüppel (2014)

Introduction

Why not use the subjective uncertainty estimates — probability bins — from SPF?

- Subjective uncertainty estimates not available from most sources of judgmental forecasts
- SPF probability forecasts are fixed event and not fixed horizon
- Flaws in SPF probability forecasts:
  - Rounding of probabilities (D'Amico & Orphanides 2008 and Boero, Smith, & Wallis 2015)
  - Overstatement of forecast uncertainty at shorter forecast horizons (Clements 2014)
  - Density forecasts from SPF histograms are no more accurate than those estimated from the historical distributions of past point forecast errors (Clements 2016)

Outline

1. Data
2. Models
3. Results
    - Full sample
    - Forecasts
4. Conclusions

Data (real-time):

Forecasts from SPF: widely studied, longest sample

- Quarterly forecasts of GDP growth, unemployment, CPI and GDP inflation, and 3-month T-bill rate
- 5 forecast horizons: $h = 0, 1, \ldots, H = 4$ quarters ahead
- A few missing obs. early in the sample
- Forecasts such as Blue Chip similar in accuracy (Reifschneider and Tulip 2007, 2017)

Data sample:

- 1969:Q1-2017:Q2: GDP growth, inflation, unemployment rate
- 1981:Q1-2017:Q2: CPI inflation, T-bill rate

Similar sample of Greenbook forecasts, through 2011

Data (real-time):

Actuals used in evaluating forecasts:

- GDP growth, GDP inflation: 1st available estimate in Phil. Fed.'s RTDSM
- Other variables: current series, from St. Louis Fed's FRED

Models

To see multi-horizon complications, consider AR-SV:

$$\begin{aligned} y_t &= b\, y_{t-1} + \sqrt{\lambda_t}\, \varepsilon_t,\ \ \varepsilon_t \sim N(0,1) \\ \log(\lambda_t) &= \log(\lambda_{t-1}) + \nu_t,\ \ \nu_t \sim N(0,\phi) \end{aligned}$$

Multi-step forecast error and error variance:

$$e_{t+h} = \lambda_{t+h}^{0.5}\epsilon_{t+h} + b\lambda_{t+h-1}^{0.5}\epsilon_{t+h-1} + \cdots + b^{h-1}\lambda_{t+1}^{0.5}\epsilon_{t+1},$$

$$\text{Var}_t\, y_{t+h} = \boldsymbol{\lambda_t} \sum_{j=0}^{h-1} b^{2j} \exp\left(\frac{1}{2}(h-j)\phi\right)$$

- Everything determined from single univariate processes
- e_{t+h} is serially correlated (i.e., correlated across h)
- $\text{Var}_t\, y_{t+h}$ is perfectly correlated across h

Models

Information set of average SPF respondent: Ω_t

- ${}_t y_{t+h} = E\left(y_{t+h} | \Omega_t\right)$
- Ω_t spans public information through $t-1$
- y_t not spanned by Ω_t

Information available from SPF at each t, for each variable y:

- Forecasts ${}_t y_{t+h}$, $h = 0, \ldots, H$, $H = 4$
- We don't know how forecasts are constructed; we take the forecasts and forecast errors as primitives
- Historical forecast errors, ${}_t e_{t+h}$, $h = 0, \ldots, H$
- ${}_t e_t$ = nowcast error

Models

Consider expectational updates:

- $\mu_{t+h|t} \equiv {}_{t}y_{t+h} - {}_{t-1}y_{t+h} = (E_t - E_{t-1})y_{t+h}$: update of forecast for $t+h$ from period $t-1$ to period t
- $\mu_{t+h|t}$ is MDS: $E_{t-1}\mu_{t+h|t} = E_{t-1}(E_t - E_{t-1})y_{t+h} = 0$

Models

Forecast error accounting identity:

$$
\begin{aligned}
{}_{t}e_{t} &\equiv y_t - E_t y_t \\
{}_{t}e_{t+1} &\equiv y_{t+1} - E_t y_{t+1} \\
&= (y_{t+1} - E_{t+1} y_{t+1}) + (E_{t+1} - E_t) y_{t+1} \\
{}_{t}e_{t+h} &\equiv (y_{t+h} - E_{t+h} y_{t+h}) + \sum_{i=1}^{h} (E_{t+h} - E_{t+h-1}) y_{t+h} \\
&= {}_{t+h}e_{t+h} + \sum_{i=1}^{h} \mu_{t+h|t+i}
\end{aligned}
$$

- Nowcast error reflects the information structure of the real-time forecasts; it would not appear in a simple time-series model setup

Models

Data vector of model:

$$\eta_t = \begin{bmatrix} y_{t-1} - E_{t-1}y_{t-1} \\ (E_t - E_{t-1})y_t \\ (E_t - E_{t-1})y_{t+1} \\ \vdots \\ (E_t - E_{t-1})y_{t+H-1} \end{bmatrix} = \begin{bmatrix} {}_{t-1}e_{t-1} \\ \mu_{t|t} \\ \mu_{t+1|t} \\ \vdots \\ \mu_{t+H-1|t} \end{bmatrix}$$

Forecast errors are linear combinations of η_{t+h}:

$$e_t = \begin{bmatrix} {}_te_t \\ \vdots \\ {}_{t-h}e_t \end{bmatrix} = B(L)\eta_{t+1} \quad \text{where } B(L) \text{ known.}$$

Models

Key: Use of $\mu_{t+h|t}$ eliminates serial correlation; η_t is an MDS

$$\mu_{t+h|t} = (E_t - E_{t-1})y_{t+h}$$
$$\Rightarrow \quad E_{t-1}\eta_t = 0$$

Key implication of treating
survey forecasts as rational expectations

Models

Multivariate stochastic volatility specification:

$$\eta_t = A \Lambda_t^{0.5} \epsilon_t \qquad A = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ a_{21} & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ a_{H+1,1} & a_{H+1,2} & \dots & & 1 \end{bmatrix}$$

$$\Lambda_t \equiv \text{diag}(\lambda_{1,t}, \dots, \lambda_{H+1,t}), \; \epsilon_t \sim N(0, I_{H+1}),$$

$$\log(\lambda_{i,t}) = \log(\lambda_{i,t-1}) + \nu_{i,t}, \quad i = 1, \dots, H+1,$$

$$\nu_t \equiv (\nu_{1,t}, \nu_{2,t}, \dots, \nu_{H+1,t})' \sim N(0, \Phi).$$

- $\text{Var}(\eta_t) = A \Lambda_t A'$
- A and Φ permit correlations of η levels and volatilities
- For forecasts from a simple time series model, the components of η would be perfectly correlated

Models

Some studies find bias and information rigidities in survey forecasts

- In our data, BIC suggests 0-1 lags for VAR in η_t
- To allow for possible biases and persistence in forecast errors and expectational updates, we extend the model to allow VAR dynamics (i.e., to not impose the MDS assumption)

Model extended to allow VAR dynamics:

$$\eta_t = C_0 + C_1 \eta_{t-1} + A {\Lambda_t}^{0.5} \epsilon_t$$

Models

Estimate the models by Bayesian MCMC methods for multivariate SV

- Gibbs sampler as in Primiceri's (2005) implementation of Kim, Shephard, and Chib (1998)
- Modified to allow for some missing values
- Priors range from uninformative to modestly informative

Simulate the posterior distribution of forecast errors

- Simulate volatility processes forward
- Simulate innovations to η forward
- Form sums according to the accounting decomposition to get back draws of the forecast errors for each horizon h
- From the posterior distribution, compute objects of interest: confidence intervals, density scores, etc.

Models

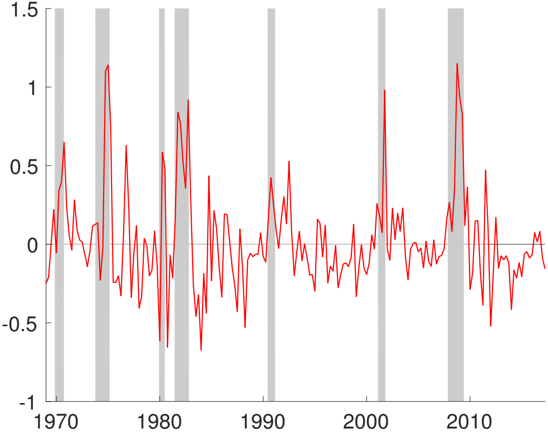
Constant forecast error variance for comparison

$$E_t e_{t+h} \sim N(0, \sigma_h^2)$$

- Similar to approach of Reifschneider and Tulip (2007, 2017)
- Applied directly to observed forecast error history
- σ_h^2 given by MSE over previous 60 quarters
- Estimated separately across h

Results: full sample

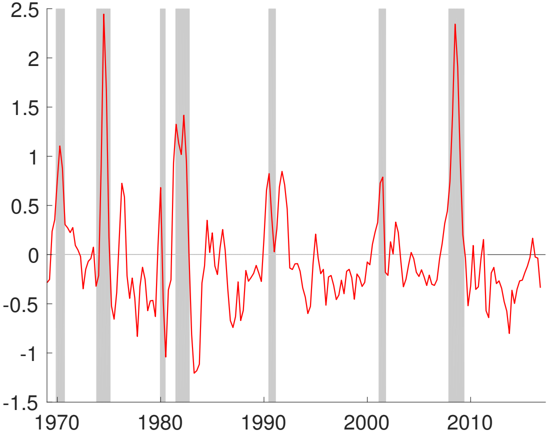
Unemployment rate, $h = 2$, red: η_t



- Expectational updates noisy

Results: full sample

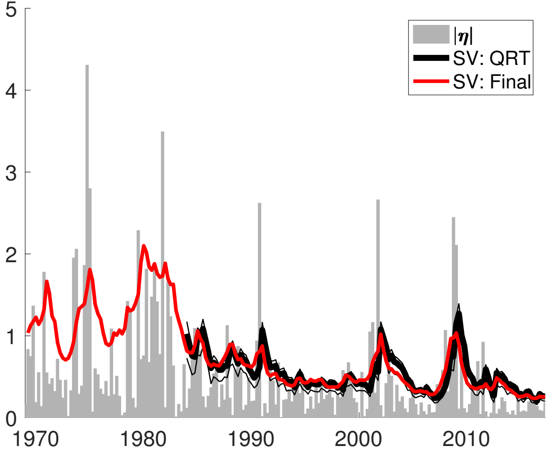
Unemployment rate, $h = 2$, red: forecast error



- Compared to updates, forecast errors are larger and more serially correlated

Results: full sample

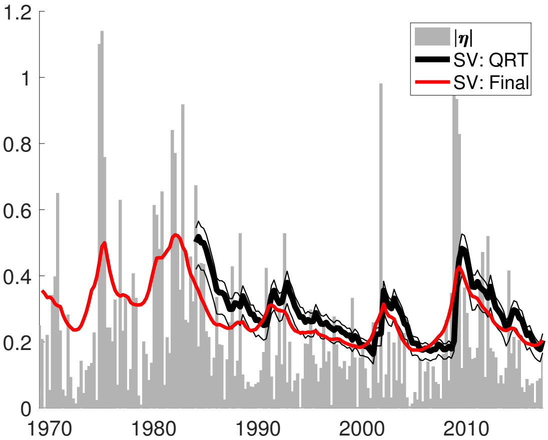
SV in η for GDP growth, $\boldsymbol{h = 2}$



- Sizable variation in volatility: Great Moderation and around recessions

Results: full sample

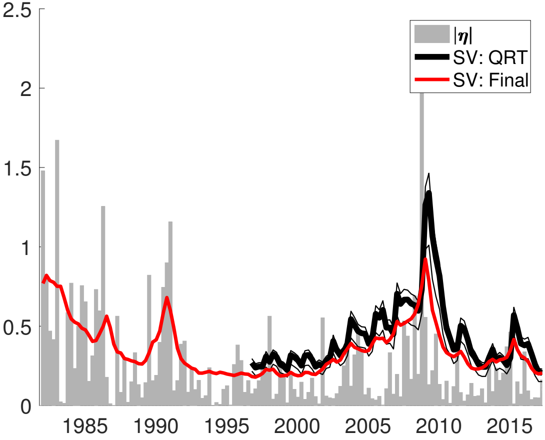
SV in η for unemployment rate, $\boldsymbol{h} = 2$



- QRT similar to ex post, but with some delay

Results: full sample

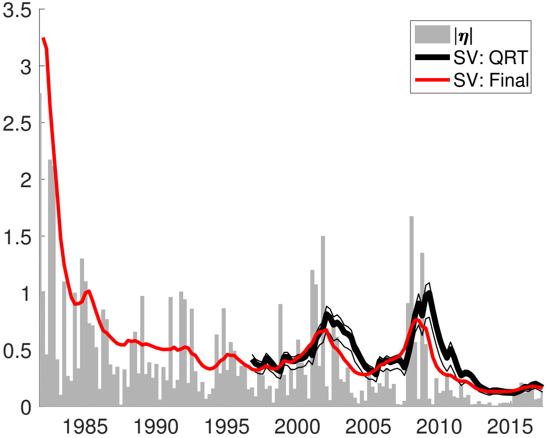
SV in η for CPI inflation, $\boldsymbol{h = 2}$



- Commodities-related spike in CPI volatility in Great Recession

Results: full sample

SV in η for T-bill rate, $\boldsymbol{h = 2}$



Results: out-of-sample forecasts

SV model

For every $t > 60$:

- Estimate model with SV using data on η_t through $t-1$
- Forecast $\text{Var}_{t-1}(\eta_{t+h})$
- Construct $\text{Var}_{t-1}(e_{t+h})$

FE-CONST approach

For every $t > 60$:

- Using forecast errors, compute $\sigma_h^2 = \text{MSE}$ for last 60 quarters
- Model predictive density with $e_{t+h} \sim N(0, \sigma_h^2)$

Results: out-of-sample forecasts

Evaluation metrics

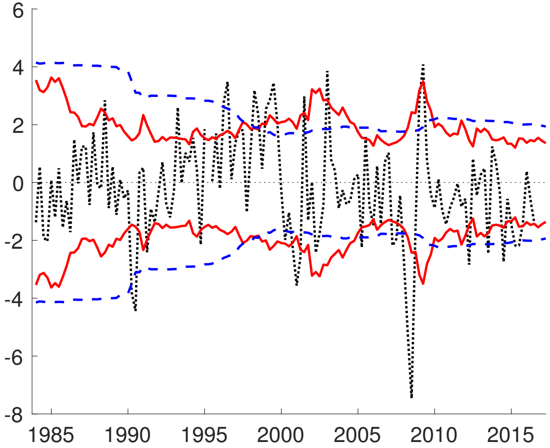
Compare SV against CONST based on

- Interval forecasts:
  - Coverage rates of one-standard-deviation bands (68%)
- Density forecast accuracy: Continuous ranked probability score (CRPS)

$$
\begin{aligned}
CRPS_t(y^o_{t+h}) &= \int_{-\infty}^{\infty} \left(F(z) - 1\{y^o_{t+h} \leq z\}\right)^2 dz \\
&= E_f|Y_{t+h} - y^o_{t+h}| - 0.5E_f|Y_{t+h} - Y'_{t+h}|
\end{aligned}
$$

Results: out-of-sample forecasts

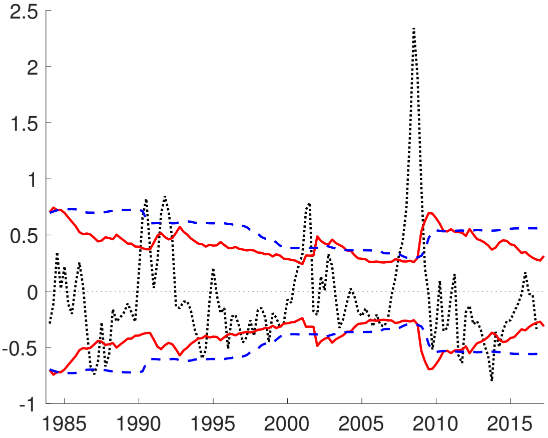
Uncertainty bands and forecast errors, GDP growth, $h = 2$



- Considerable time variation in band widths, more so with SV
- For much of the sample SV band narrower than CONST band

Results: out-of-sample forecasts

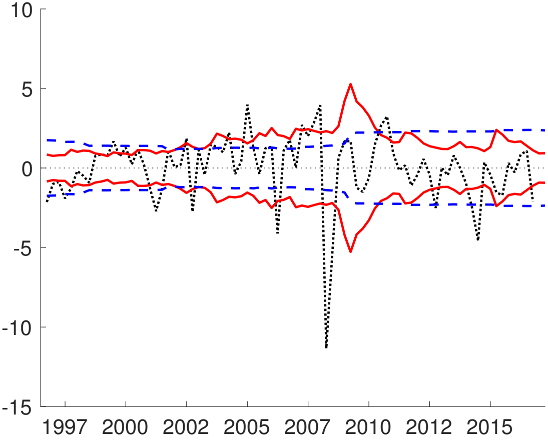
Uncertainty bands and forecast errors, unemployment, $h = 2$



- Crisis widens bands, more so for SV (temporarily) than CONST

Results: out-of-sample forecasts

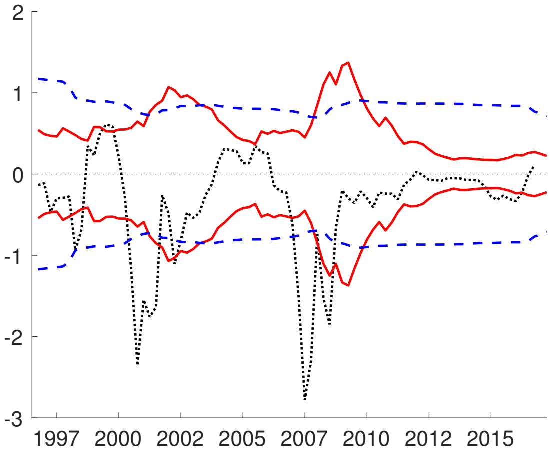
Uncertainty bands and forecast errors, CPI inflation, $\boldsymbol{h = 2}$



- CPI results different: bands widen over the sample, and SV bands wider than CONST bands

Results: out-of-sample forecasts

Uncertainty bands and forecast errors, T-bill rate, $h = 2$



Results: out-of-sample forecasts

coverage rates of one-standard deviation bands

Variable	**Forecast horizon** 0	1	2	3	4
Panel A: SV					
RGDP	72.73	71.76	72.31	72.87	69.53
UNRATE	74.63	75.19	69.70	66.41	63.85
PGDP	73.48	72.52	73.85	71.32	71.09
CPI	68.67	64.63	64.20	67.50	69.62
TBILL	80.72**	80.49**	72.84	65.00	51.90
Panel B: FE-CONST					
RGDP	76.52**	78.63**	76.92*	78.29*	79.69**
UNRATE	73.13	82.71***	87.12***	87.79***	86.92***
PGDP	74.24	78.63***	77.69**	79.07**	79.69**
CPI	71.08	63.41	67.90	66.25	70.89
TBILL	79.52**	87.80***	83.95**	80.00	78.48

- Intervals more accurate with SV than FE-CONST specification
- (evidenced in counts of significant departures from correct coverage)

Results: out-of-sample forecasts

CRPS: Percentage improvement of SV over CONST

	Forecast Horizon				
Variable	0	1	2	3	4
RGDP	3.04**	7.19***	7.55***	8.52***	6.33**
UNRATE	0.91	1.75*	2.48*	2.51	1.56
PGDP	0.58	1.61	2.37*	2.43	3.26
CPI	1.08	1.14	1.53	2.65	2.12
TBILL	8.65***	12.09***	11.20***	8.07*	5.19

- SV consistently improves on density accuracy of FE-CONST
- Gains largest for T-bill rate and GDP
- Note: gains entirely driven by uncertainty estimates

Conclusions

Our contributions:

- We derive a multi-horizon SV framework
- Bayesian estimation with MCMC/Gibbs sampler
- Document time-varying uncertainty in SPF and Greenbook forecasts

Comparing SV against rolling-window FE-CONST:

- More accurate confidence intervals (fan charts)
- More accurate densities as measured by CRPS
- Departing from MDS assumption and allowing VAR dynamics helps for some variables and not others