Point, Interval, and Density Forecast Evaluation of Linear versus Nonlinear DSGE Models

Francis X. Diebold          Frank Schorfheide          Minchul Shin

University of Pennsylvania

May 4, 2014
Motivation

The use of DSGE models as a forecasting tool in practice calls for evaluation of its performance.

Summary of DSGE forecasting evaluation literature:

- Point forecasts are as good as other statistical models such as VAR.
- Density forecasts are too wide.

Typically, DSGE models in this literature are

- Solved with linearized method.
- Constant volatility.

Forecasting with a nonlinear DSGE model:

- Pichler (2008) considers the second-order perturbation method.
Nonlinearity and Forecasting

Not much work has been done for nonlinear DSGE forecasting.

Two types of nonlinearities

- **Time-varying volatility** vs. Constant volatility
  - There are many papers on DSGE models with time-varying volatility but there is no paper that systematically evaluates forecasts.
  - Most papers report that density forecasts generated from the DSGE models with constant volatility is too wide.
- **Nonlinear approximation** vs. Linear approximation
  - Pichler (2008)’s point forecast evaluation.
  - Interval and density forecasts evaluation.
Summary

Method

- Small-scale DSGE model.
  - Linear DSGE model with constant volatility.
  - Linear DSGE model with time-varying volatility.
  - Quadratic DSGE model (based on the second-order perturbation method).
  - Quadratic DSGE model with time-varying volatility.
- Bayesian estimation and forecasting. (US data, 1964-2011)

Finding

- Modelling time-varying volatility improves accuracy of forecasts.
- Second-order perturbation method does not improve.
Model, Estimation, and Forecasting
Small-Scale DSGE Model

Small-scale model used in Herbst and Schorfheide (2012)
- Euler equation, NK Phillips curve, Monetary policy rule
- 3 exogenous shocks: technology, government spending, monetary policy \((z_t, g_t, mp_t)\)

Measurement equations:

\[
\begin{pmatrix}
YGR_t \\
INF_t \\
FFR_t
\end{pmatrix} = D(\theta) + Z(\theta) \, s_t
\]

Transition equations:

\[
s_t = \Phi(s_{t-1}, \epsilon_t; \theta)
\]

where

\[
s_t = [y_t, y_{t-1}, c_t, \pi_t, R_t, mp_t, z_t, g_t]'
\]

\(\epsilon_t\) : Innovations
\(\theta\) : DSGE parameters
Linear DSGE Model

Linear approximation methods lead to a linear Gaussian state space representation with the following transition equation

\[ s_t = H(\theta)s_{t-1} + R(\theta)\epsilon_t, \quad \epsilon_t \sim iid\mathcal{N}(0, Q(\theta)). \]

- Coefficient matrices \((H(\theta), R(\theta), Q(\theta))\) are the nonlinear function of \(\theta\).
- We obtain posterior draws based on the Random Walk Metropolis (RWM) algorithm with the Kalman filter.
Linear+Stochastic Volatility

Following Justiniano and Primiceri (2008),

\[ s_t = H(\theta)s_{t-1} + R(\theta)\epsilon_t, \quad \epsilon_t \sim \mathcal{N}(0, Q_t(\theta)) \]

where

\[
\text{diag}(Q_t(\theta)) = \left[ e^{2h_{mp,t}}, e^{2h_{z,t}}, e^{2h_{g,t}} \right]'
\]

\[ h_{i,t} = \rho_{\sigma_i} h_{i,t-1} + \nu_{i,t}, \quad \nu_{i,t} \sim \text{iid}\mathcal{N}(0, s_{i}^2), \]

for \( i = mp, g, z \).

- The system is in a linear Gaussian state-space form conditional on \( Q_t(\theta) \).
- We use the Metropolis-within-Gibbs algorithm developed by Kim, Shephard, and Chib (1998) to generate draws from the posterior distribution.
Quadratic DSGE Model

The equilibrium law of motion based on the second-order perturbation method has the following form:

\[ s_{1,t} = G_0(\theta) + G_1(\theta)s_{2,t} + G_2(\theta)(s_{2,t} \otimes s_{2,t}) \]
\[ s_{2,t} = H_0(\theta) + H_1(\theta)s_{2,t-1} + H_2(\theta)(s_{2,t-1} \otimes s_{2,t-1}) + R(\theta)\epsilon_t \]

where \( \epsilon_t \sim iid \mathcal{N}(0, Q(\theta)) \), \( s_t = [s_{1,t}, s_{2,t}]' \) and \( \otimes \) is a Kronecker product.

- We have an additional quadratic term.
- We need a nonlinear filter to get posterior draws.
  - Particle filter + RWM (exact).
  - Second-order Extended Kalman Filter + RWM (approximate).
- Current version of the paper utilizes the second-order extended Kalman filter and RWM.
- We also consider Quadratic DSGE model with stochastic volatility.
Predictive Distribution

We generate draws from the posterior predictive distribution based on the following decomposition,

\[
p(Y_{T+1:T+H}|Y_{1:T}) = \int p(Y_{T+1:T+H}|\theta, Y_{1:T}) p(\theta|Y_{1:T}) d\theta
\]

- \(p(\theta|Y_{1:T})\) : Posterior sampler.
- \(p(Y_{T+1:T+H}|\theta, Y_{1:T})\) : Given \(\theta\), simulate the model economy forward.

Draws \(\{Y_{T+1:T+H}^{(j)}\}_{j=1}^{n_{sim}}\) can be turned into point and interval forecasts by the Monte Carlo average,

\[
E[y_{T+h}|T] = \int_{y_{T+h}} y_{T+h} p(y_{T+h}|Y_{1:T}) dy_{T+h} \approx \frac{1}{n_{sim}} \sum_{j=1}^{n_{sim}} y_{T+h}^{(j)}.
\]
Evaluation Methods
Point Forecast

Point forecasts from two different models (Inflation rate)

We compare RMSEs for variable $i = \{YGR, INFL, FFR\}$,

$$RMSE(i|h) = \sqrt{\frac{1}{P-h} \sum_{t=R}^{R+P-h} (y_{i,t+h} - \hat{y}_{i,t+h|t})^2}$$

where $R$ is the index denotes the starting point of the forecasts evaluation sample and $P$ the number of forecasting origins.
We compute $\alpha\%$ interval forecasts for a particular element $y_{i,T+h}$ of $y_{T+h}$ by numerically searching for the shortest connected interval that contains a $\alpha\%$ of the draws $\left\{y_{i,T+h}^{(j)}\right\}_{j=1}^{n_{sim}}$ (the highest-density set).

If the interval forecast is well calibrated, actual variables are expected to be inside of $\alpha\%$ interval forecasts at the same frequency.
Interval Forecast

Interval forecasts from two different models (Inflation rate)

Christoffersen (1998)’s LR tests of the correct coverage ($\alpha$):

- Define the sequence of hit indicators of a 1-step-ahead forecast interval,
  
  \[ I_t^\alpha = 1\{\text{realized } y_t \text{ falls inside the interval}\} \]

- Test
  
  \[ I_t^\alpha \sim iid \text{Bernoulli}(\alpha). \]
Density Forecast

Density forecasts from two different models (Inflation rate)

The **probability integral transform (PIT)** of \( y_{i,T+h} \) based on time \( T \) predictive distribution is defined as the cumulative density of the random variable,

\[
Z_{i,h,T} = \int_{-\infty}^{y_{i,T+h}} p(\tilde{y}_{i,T+h} | Y_{1:T}) d\tilde{y}_{i,T+h}.
\]

- If the predictive distribution is well-calibrated, \( z_{i,h,T} \) should follow the uniform distribution.
- For \( h = 1 \), \( z_{i,h,T} \)’s follow independent uniformly distribution.
Density Forecast Evaluation: Predictive Likelihood

The one-step-ahead predictive likelihood,

\[ PL(t) = p(y_{t+1} | Y_{1:t}) \]

- Height of the predictive density at the realized value \( y_{t+1} \).
- Log predictive score: \( \sum_{t=R}^{R+P-1} \log PL(t) \).

We approximate

\[ PL(t) \approx \frac{1}{M} \sum_{m=1}^{M} p \left( y_{t+1} | Y_{1:t}, \tilde{\theta}^{(m)} \right) \]

where \( \{ \tilde{\theta}^{(m)} \} \) is a sequence from the posterior simulator using data \( Y_{1:t} \).
Results
Data

For the evaluation of forecasts, we use the real-time data set constructed by Del Negro and Schorfheide (2012).

- Forecast horizons and data vintages are aligned with Blue Chip survey publication dates.
- Output growth, Inflation, Federal Funds rate.
- Generate forecasts four times a year (January, April, July, and October).

To evaluate forecasts we recursively estimate DSGE models over the 78 vintages starting from January 1992 to April 2011.

- All estimation samples start in 1964.
- Compute forecast errors based on actuals that are obtained from the most recent vintage.
Data

Data, 1964Q2-2011Q1
Standard deviations of the structural shocks (Posterior mean):

- Dotted : Linear DSGE with constant volatility
- Solid : Linear DSGE with SV
Point Forecasts Evaluation: RMSEs

Linear vs Linear+SV, 1991Q4-2011Q1

- SV improves point forecasts for output growth.
- Difference in RMSEs for output growth is significant (Diebold and Mariano, 1995)
Point Forecasts Evaluation: RMSEs

Quadratic vs Quadratic+SV, 1991Q4-2011Q1

- SV improves point forecasts for output growth for Quadratic model as well.
Point Forecasts Evaluation: RMSEs

Linear+SV vs Quadratic+SV, 1991Q4-2011Q1

- Quadratic term does not improve point forecasts.
- Similar results as in Pichler (2008) but with SVs.
Interval Forecasts: Coverage Rate

Coverage Rate of 70% Interval Forecasts, 1991Q4-2011Q1

If the interval forecast has a correct coverage rate, then it should be around 0.7 line (dotted black line).

- Interval forecasts from the models with SV are closer to 0.7 line.
- Quadratic+SV has better coverage rates for inflation rate $h \geq 3$. 
Interval Forecasts

Output Growth Interval Forecasts \((h = 1)\), 1991Q4-2011Q1

- In general, interval forecasts are shorter for the model with SV.
- pre-Great moderation sample effect.
Interval Forecasts

In general, interval forecasts are shorter for the model with SV.

pre-Great moderation sample effect.

Rolling window estimation (with 80Q) helps but not much.
Density Forecasts Evaluation: PITs, \((h = 1)\)

PITs, 1-Step-Ahead Prediction, 1991Q4-2011Q1

PITs are grouped into five equally sized bids. Under a uniform distribution, each bin should contain 20% of the PITs, indicated by the solid horizontal lines in the figure.
Density Forecasts Evaluation: PITs \((h = 1)\)

PITs, 1-Step-Ahead Prediction, 1991Q4-2011Q1

**Quadratic DSGE Model**

PITs are grouped into five equally sized bids. Under a uniform distribution, each bin should contain 20% of the PITs, indicated by the solid horizontal lines in the figure.

**Quadratic + SV**

[Graphs showing PITs for GDP Growth, Infl. Rate, and Int. Rate for both models.]
Density Forecasts Evaluation: PITs \((h = 4)\)

PITs, 4-Step-Ahead Prediction, 1991Q4-2011Q1

Linear DSGE Model

Linear+SV

PITs are grouped into five equally sized bids. Under a uniform distribution, each bin should contain 20% of the PITs, indicated by the solid horizontal lines in the figure.
Density Forecasts Evaluation: PITs ($h = 4$)

PITs, 4-Step-Ahead Prediction, 1991Q4-2011Q1

Quadratic DSGE Model

Quadratic + SV

PITs are grouped into five equally sized bids. Under a uniform distribution, each bin should contain 20% of the PITs, indicated by the solid horizontal lines in the figure.
Density Forecasts Evaluation: Normalized Error

\[ w_{i,t} = \Phi^{-1}(z_{i,t}) \]

LR Tests of normalized errors of 1-step ahead real-time forecasts

<table>
<thead>
<tr>
<th></th>
<th>Std. Dev. (= 1)</th>
<th>Mean (= 0)</th>
<th>AR(1) coef. (= 0)</th>
<th>LR test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(a) Output Growth</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>0.543 (0.000)</td>
<td>0.195 (0.005)</td>
<td>0.357 (0.021)</td>
<td>53.049 (0.000)</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.547 (0.000)</td>
<td>0.206 (0.003)</td>
<td>0.364 (0.018)</td>
<td>52.984 (0.000)</td>
</tr>
<tr>
<td>Linear+SV</td>
<td><strong>0.827 (0.349)</strong></td>
<td><strong>0.093 (0.493)</strong></td>
<td>0.254 (0.072)</td>
<td>10.610 (0.014)</td>
</tr>
<tr>
<td>Quadratic+SV</td>
<td><strong>0.845 (0.422)</strong></td>
<td><strong>0.146 (0.282)</strong></td>
<td>0.315 (0.022)</td>
<td>13.547 (0.004)</td>
</tr>
<tr>
<td><strong>(b) Inflation Rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td><strong>0.762 (0.316)</strong></td>
<td><strong>0.058 (0.597)</strong></td>
<td><strong>-0.125 (0.474)</strong></td>
<td>10.595 (0.014)</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.738 (0.045)</td>
<td>0.075 (0.598)</td>
<td>0.119 (0.273)</td>
<td>12.694 (0.005)</td>
</tr>
<tr>
<td>Linear+SV</td>
<td><strong>0.868 (0.461)</strong></td>
<td><strong>0.130 (0.368)</strong></td>
<td><strong>-0.000 (0.998)</strong></td>
<td>3.915 (0.271)</td>
</tr>
<tr>
<td>Quadratic+SV</td>
<td><strong>0.890 (0.592)</strong></td>
<td><strong>0.050 (0.777)</strong></td>
<td><strong>0.090 (0.516)</strong></td>
<td>2.569 (0.463)</td>
</tr>
<tr>
<td><strong>(c) Fed Funds Rate</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear</td>
<td>0.578 (0.000)</td>
<td>-0.070 (0.658)</td>
<td>0.663 (0.000)</td>
<td>77.837 (0.000)</td>
</tr>
<tr>
<td>Quadratic</td>
<td>0.615 (0.000)</td>
<td>-0.067 (0.705)</td>
<td>0.672 (0.000)</td>
<td>73.064 (0.000)</td>
</tr>
<tr>
<td>Linear+SV</td>
<td><strong>0.866 (0.492)</strong></td>
<td><strong>-0.119 (0.593)</strong></td>
<td>0.744 (0.000)</td>
<td>66.514 (0.000)</td>
</tr>
<tr>
<td>Quadratic+SV</td>
<td><strong>0.859 (0.483)</strong></td>
<td><strong>-0.073 (0.743)</strong></td>
<td>0.712 (0.000)</td>
<td>58.611 (0.000)</td>
</tr>
</tbody>
</table>
Log Predictive Density

<table>
<thead>
<tr>
<th></th>
<th>$h = 1Q$</th>
<th>$h = 2Q$</th>
<th>$h = 4Q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>-3.99</td>
<td>-4.20</td>
<td>-4.91</td>
</tr>
<tr>
<td>Quadratic</td>
<td>-3.97</td>
<td>-4.49</td>
<td>-5.05</td>
</tr>
<tr>
<td>Linear+SV</td>
<td><strong>-3.82</strong></td>
<td>-4.66</td>
<td>-5.70</td>
</tr>
<tr>
<td>Quadratic+SV</td>
<td>-3.84</td>
<td>-4.63</td>
<td>-5.52</td>
</tr>
</tbody>
</table>

- Linear+SV performs the best for the 1-step-ahead prediction.
- Linear without SV performs better for $h \geq 2$. 
Conclusion

Summary

- Modelling the time variation improves the density forecasts especially in the short-run.
- The second order perturbation method does not improve forecasts’ quality in general.

Future works

- More features
  - Time-varying inflation target.
  - SV process: ARMA(p,q) as opposed to AR(1).
- Larger model
- Particle filter