# *Forecasting*

May 23, 2011

# Today:

Forecasting

- Overview
- Forecast combination
- Density forecasts

Based on Stock and Watson (2005) and Gerard and Nimark (2009)

# Forecasting

What is it?

Producing a prediction about the value of a random variable $y$ at horizon $h$ conditional on the information set in period $t$

$$y_{t+h|t} = f(\Omega_t)$$

Sometimes one want to predict the distribution so that the object of interest is

$$p(y_{t+h|t}) = \widetilde{f}(\Omega_t)$$

## What are we forecasting?

Generally, the future. But there are variants of what the "future" means:

- In sample forecasting (forecasts use what they forecast to estimate parameters of model)
- Out of sample forecasting
- Pseudo out-of-sample forecasting

True out-of-sample forecasts have to be collected from something like a public record

# Pseudo out-of-sample forecasting

Divide sample into training sample and forecasting sample

- Use training sample to estimate model used to forecast "out-of-sample"

Examples of strategies:

- Expanding window: Estimate model on sample running from $1, 2, ..., t$ and use to produce forecasts of variables at date $t + h : h > 0$;
- Rolling window: Estimate model on sample running from $t - s, t - s + 1, ...t$ and use to produce forecasts of variables at date $t + h : h > 0$;

Rolling window often used when parameter instability or structural change is suspected

# Alternative specifications

Iterated one-step-ahead forecasts vs $h$ step ahead forecasts

$$y_{t+h} = A_h y_t + \varepsilon_t$$

or

$$\begin{aligned} y_{t+h} &= A_1^h y_t \\ y_t &= A_1 y_{t-1} + \varepsilon_t \end{aligned}$$

No clear consensus which approach is better

- With correctly specified model, $h$ iterated 1-step-ahead forecasts should be better since more efficient estimates of parameters

# Producer and consumers of forecasts

Why does it matter what you will do with a forecast?

There may be biased forecasts with small MSE and unbiased forecasts with large MSE.

- Two different users may disagree which forecast is best for them

There is a large decision theoretic literature linking utility functions and optimality of forecast

# Forecasting in a data rich environment

Macro economists have a peculiar data situation:

- Many data series, but usually short samples

How can we utilize all this information without running into degrees of freedom problems?

# Pitfalls of OLS when $N$ is large relative to $T$

Degrees of freedom problems

- You cannot have $N$ larger than $T$ and run a VAR

It can be shown that if $N$ grows proportionally to $T$, standard errors of parameter estimates do not shrink with sample size

- Forecast errors can be larger than the unconditional error(!)

So, what can be done?

# A general principle: Dimension reduction

Dimension reduction usually improves forecast accuracy

- Simple models often do better. A univariate AR(1) is hard to beat...
- BVARs (w/Litterman prior): A combination of a univariate unit roots and a VAR
- Factor models
- DSGE models
  - Using economic theory to restrict the forecasting model

# Forecast combination

How can we deal with model uncertainty?

Forecast combination:

- Can work well and forecast combinations often outperform all individual forecasts used in pool

What is it?

$$y_{t+h|t} = \sum w_i y^i_{t+h|t}$$

The set $y^i_{t+h|t}$ is also known as the "forecast pool"

# Forecast combination

How can we choose the weights $w_i$?

- Weights based on in or out of sample fit measures
  - In-sample $R^2$
  - Previous out-of-sample fit

In practice, this adds additional free variables which may be undesirable

# How can we control for "similar" models?

If there are many similar models we do not get the full benefit of "model diversification"

- Choose weights minimize Kullback-Leibler distance between forecast and "truth"
  - We don't know the "true" model, but by definition, outcomes are generated by the "true" model

For point forecasts, this simply means regressing previous forecasts on outcomes to find weights:

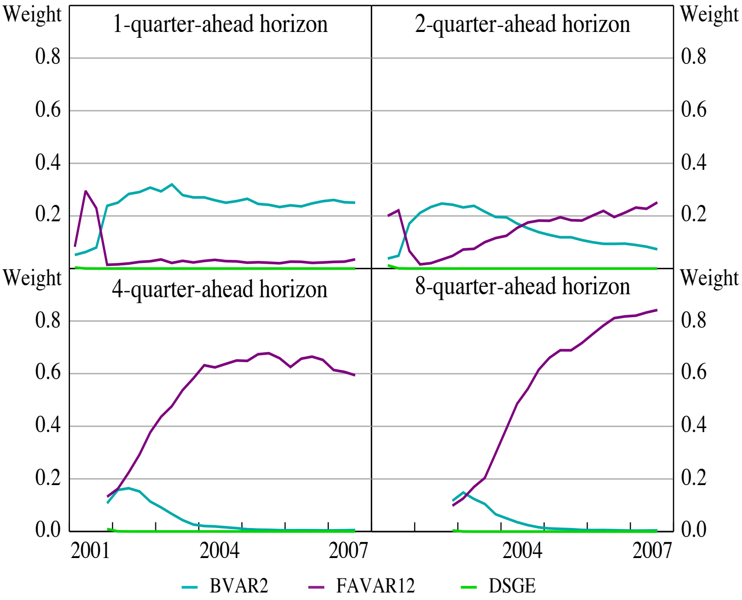
$$y_{t+h} = \sum w_i y^i_{t+h|t} + u_t$$

But simple combinations like equal weights on all models often work better than sophisticated methods (also known as the "forecast combination puzzle").

# Forecast combination

Sometimes the models in a forecast pool that are good at short horizons are not so good at long horizon forecasting

- This can be accommodated by choosing horizon specific weights



Weight
1-quarter-ahead horizon
2-quarter-ahead horizon
4-quarter-ahead horizon
8-quarter-ahead horizon
0.8
0.6
0.4
0.2
0.0
2001
2004
2007
BVAR2
FAVAR12
DSGE

## Forecast combination

Few general lessons:

- Why does combinations work so well?
- Why do simple combinations often do as well as more sophisticated combinations?

Probably it is because none of the models we can think of is really the true data generating process.

# Other types of forecasts

Density forecasts (also known as fan charts)

- Provides a more complete picture of the distribution of future events than a simple point forecast
- Communicates uncertainty around forecasts

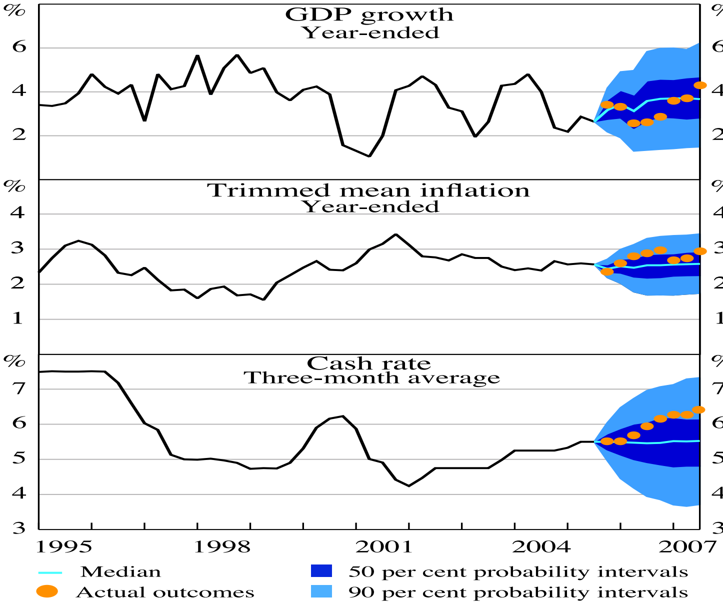
Popular with central banks nowadays

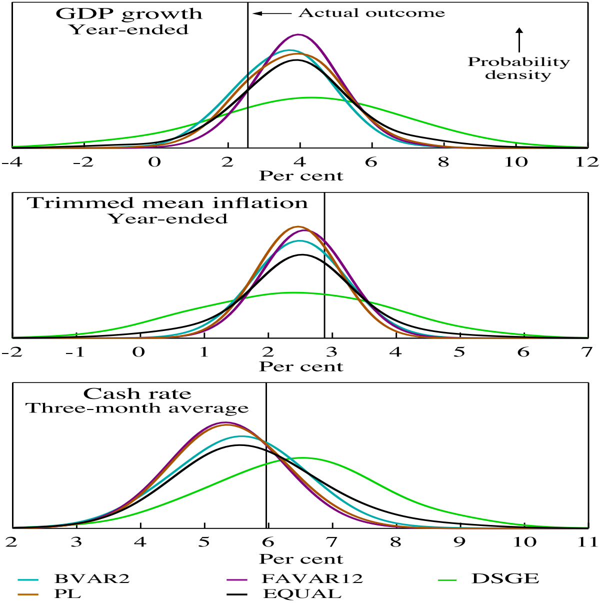
# Density forecast example

Inflation targeting central bank

- What is the probability that the inflation rate will be outside the official target band?

Have become popular tools to communicate interest rate expectations to the public without appearing to "commit" to a particular path


GDP growth
Year-ended
Trimmed mean inflation
Year-ended
Cash rate
Three-month average
%
1995
1998
2001
2004
2007
Median
Actual outcomes
50 per cent probability intervals
90 per cent probability intervals


GDP growth
Year-ended
Actual outcome
Probability density
-4 -2 0 2 4 6 8 10 12
Per cent
Trimmed mean inflation
Year-ended
-2 -1 0 1 2 3 4 5 6 7
Per cent
Cash rate
Three-month average
2 3 4 5 6 7 8 9 10 11
Per cent
BVAR2
PL
FAVAR12
EQUAL
DSGE

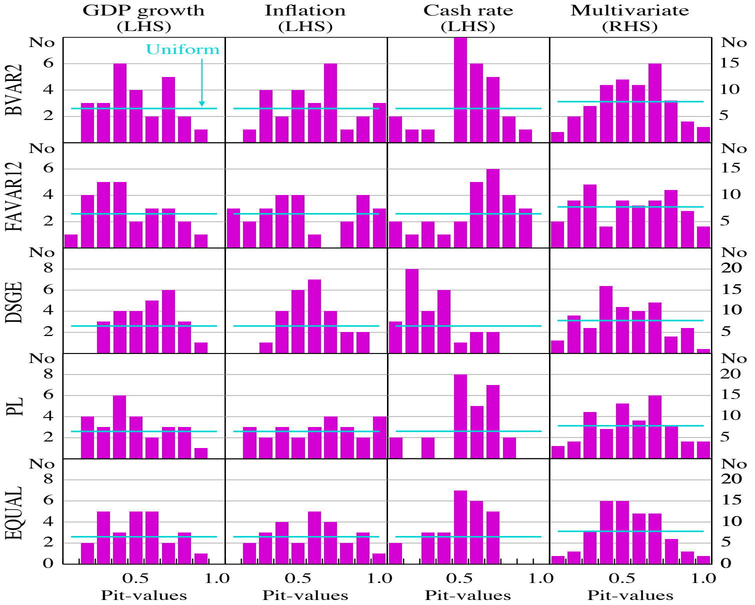
# How can density forecasts be evaluated?

What does precision mean for density forecasts?

- Point forecasts should be close to the actual outcomes

One approach is to use Probability Integral Transforms (PITs)

- Create "bins" for each decile and count frequency for each bin (should be 10%)



GDP growth (LHS)
Inflation (LHS)
Cash rate (LHS)
Multivariate (RHS)
BVAR2
FAVAR12
DSGE
PL
EQUAL
Uniform
No
Pit-values
0.5
1.0

# Forecasting: Summing up

Forecasting is perhaps the most common usage of time series models in practice

- Which models are good or bad may depend on who will use it

Dimension reduction is often important

- Avoids in-sample over fitting

Point- and density forecast require different evaluation tools

- PITs is one way, but shortness of sample can be real problem