Discussion of "Financial Conditions and Density Forecasts for US Output and Inflation" by Piergiorgio Alessandri (Bank of Italy) with Haroon Mumtaz (Queen Mary)

Shaun Vahey (Warwick)

8^{th} Workshop on Forecasting Techniques, ECB

June 2014

Questions from the paper

- Do financial conditions (measured by fci) improve forecast accuracy for US inflation, output (and other variables)?
- Do threshold VARs perform better?
- Did financial conditions provide a “credible warning” (in real time) of Great Recession? Raising normative issues ...

Questions from the paper

- Do financial conditions (measured by fci) improve forecast accuracy for US inflation, output (and other variables)?
- Do threshold VARs perform better?
- Did financial conditions provide a "credible warning" (in real time) of Great Recession? Raising normative issues ...
- Main results: Yes, Yes and Yes (please, next time policymakers)

Some questions of interest (to me))

- Can log score based evaluations of forecast densities mask predictive content?
- What additional steps—beyond log scores—might be useful to analyse forecast performance?
- What characteristics do policymakers want from a forecast when seeking advanced warnings?

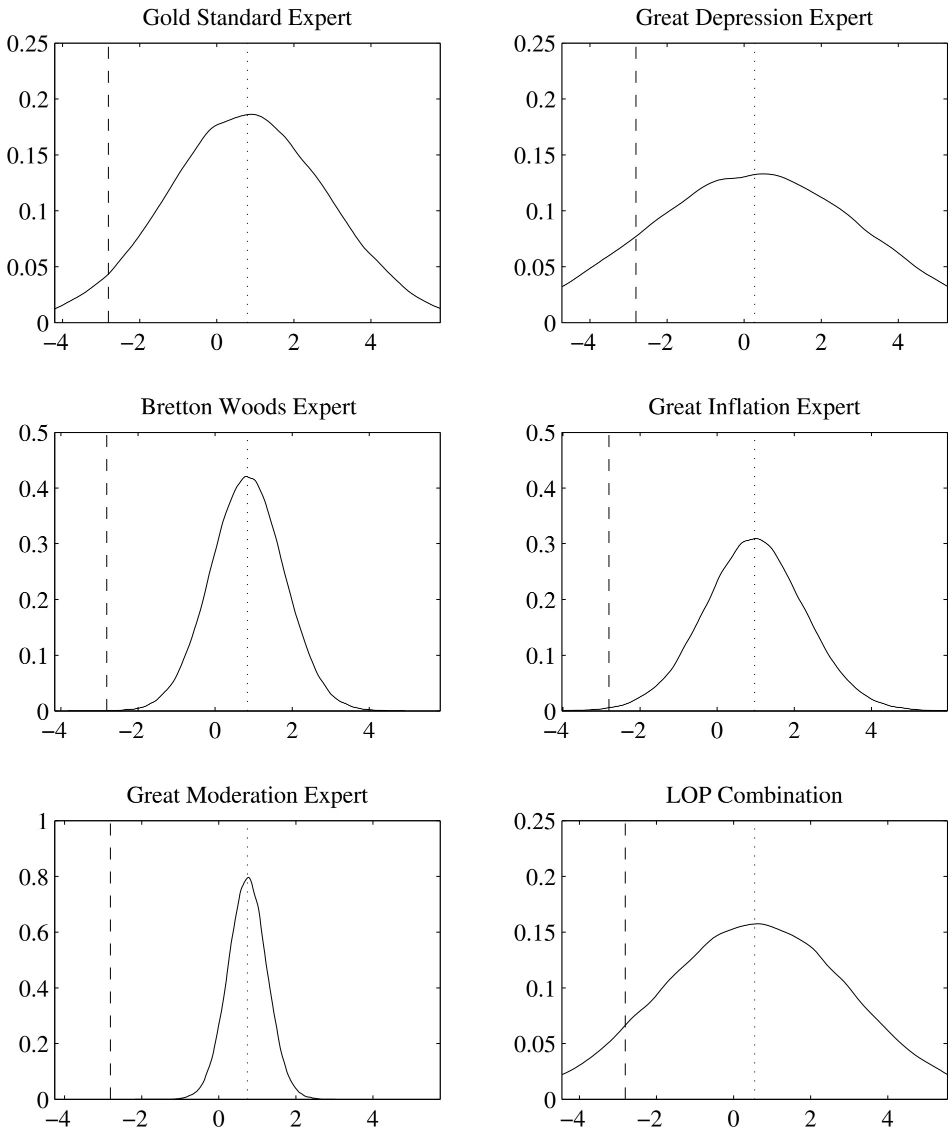
An example to illustrate some issues

- "Experts", Great Depression era, and Bretton Woods era, plus a few friends; each uses a bivariate VAR in inflation and output estimated on data only from relevant era
- Then, look at a bake-off between the two experts through Great Recession, and consider RMSFE, log scores
- In this example, despite a strong log score performance, the policymaker wouldn't want to bring back a defunct Great Depression expert to call the slump ...

Table 4: Log Scores Relative to Great Moderation Expert

	2005Q1-2010Q4	2005Q1-06Q4	2007Q1-08Q4	2009Q1-10Q4
(a) Output Growth				
Gold Standard	0.736*	0.833	0.600	0.841
Great Depression	0.723*	0.829	0.576	0.832
Bretton Woods	0.839*	0.872	0.770	0.907
Great Inflation	0.855*	0.903	0.743	0.973
Great Moderation	1.000	1.000	1.000	1.000
Linear Opinion Pool	0.726*	0.827	0.585	0.831

Figure 3: Output Growth Forecast Densities 2008Q4



Notes: The vertical dashed lines indicate the location of the realization in 2008Q4 and the vertical dotted lines indicate the means of the forecast densities.

Cost-loss approach with negative output growth events

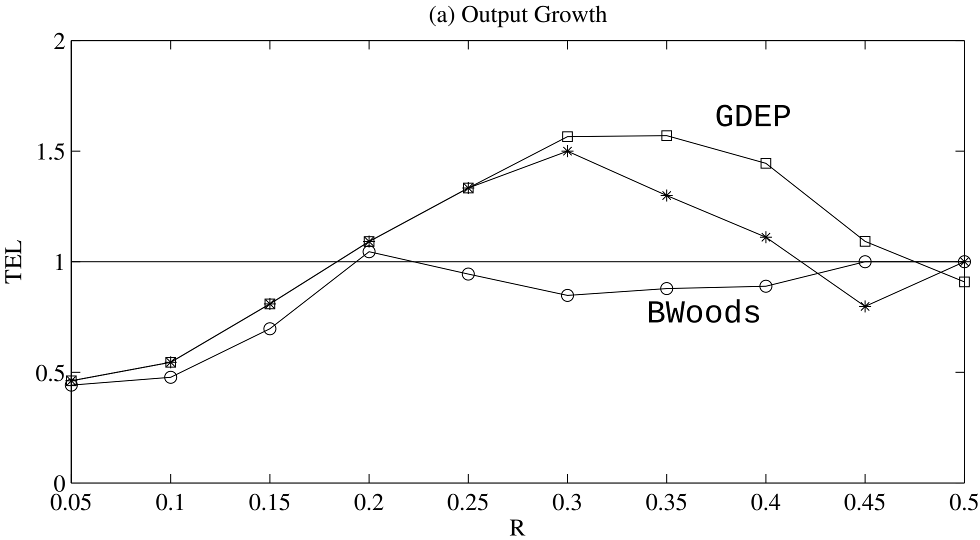
- Following eg Granger and Pesaran (2000), Berrocal et al (2010), relative cost of unanticipated contraction $R = C/L$, $0 < R < 1$, unknown

- Issue contraction warning only if $Pr(\Delta y < 0) > R$

- Define $TEL = n_{10}L \; + \; (n_{01} + n_{00})C$

	Event Observed	
Event Forecast	Yes	No
Yes	n_{00}	n_{01}
No	n_{10}	n_{11}

Figure 7: Economic Loss Relative to Great Moderation Expert



Back to those questions of interest (to me))

- Can log score based evaluations of forecast densities mask predictive content?

Back to those questions of interest (to me))

- Can log score based evaluations of forecast densities mask predictive content?

- Yep. Strong relative log scores differentials aren't sufficient to indicate that the policymaker can use the model (expert) in real time to give an early warning indicator

Back to those questions of interest (to me))

- What additional steps—beyond log scores—might be useful to analyse forecast performance? Some tricks from the (forthcoming) PROFOR toolbox (spvahey@gmail.com)

Back to those questions of interest (to me))

- What additional steps—beyond log scores—might be useful to analyse forecast performance? Some tricks from the (forthcoming) PROFOR toolbox (spvahey@gmail.com)

    1. CRPS Hersbach (2000), Ravazzolo-Vahey (2009, 2013) maximising sharpness conditional on calibration; plus threshold scoring rules, Gneiting-Ranjan (2011), Garratt-Mitchell-Vahey (2013); compare with common benchmark; market timing statistics, Pesaran-Timmermann (1992)

Back to those questions of interest (to me))

- What additional steps—beyond log scores—might be useful to analyse forecast performance? Some tricks from the (forthcoming) PROFOR toolbox (spvahey@gmail.com)

    1. CRPS Hersbach (2000), Ravazzolo-Vahey (2009, 2013) maximising sharpness conditional on calibration; plus threshold scoring rules, Gneiting-Ranjan (2011), Garratt-Mitchell-Vahey (2013); compare with common benchmark; market timing statistics, Pesaran-Timmermann (1992)
    2. Analyse the PITS to check calibration; eg Diebold-Gunther-Tay (1998), Jore-Mitchell-Vahey (2010)

Back to those questions of interest (to me))

- What additional steps—beyond log scores—might be useful to analyse forecast performance? Some tricks from the (forthcoming) PROFOR toolbox (spvahey@gmail.com)

    1. CRPS Hersbach (2000), Ravazzolo-Vahey (2009, 2013) maximising sharpness conditional on calibration; plus threshold scoring rules, Gneiting-Ranjan (2011), Garratt-Mitchell-Vahey (2013); compare with common benchmark; market timing statistics, Pesaran-Timmermann (1992)
    2. Analyse the PITS to check calibration; eg Diebold-Gunther-Tay (1998), Jore-Mitchell-Vahey (2010)
    3. Plot forecast densities and check out the shape (various moments could be "post-processed")

Back to those questions of interest (to me))

- What characteristics do policymakers want from a forecast when seeking advanced warnings?

Back to those questions of interest (to me))

- What characteristics do policymakers want from a forecast when seeking advanced warnings?
- Utilise loss function, Granger-Pesaran (2000a, 2000b), Coe-Vahey (2014)
- With unknown loss function, need to check out calibration (reliability), as noted by Jore-Mitchell-Vahey (2010), plus resolution vNorden-Galbraith (2008) and/or sharpness

In summary

- An interesting paper and an important issue
- Perhaps a little more to do in describing and understanding the differences in forecast performance

Surely, density forecasting and loss-based evaluation is a promising route to deal with the quacks ...?

- John Kay (FT, September 21 2010):

  "There will always be a demand for forecasts, so there will always be a supply. But the reputation of economic forecasters, like other quacks and charlatans, depends more on the slickness of their presentations than the value of their work"