Discussion of the paper: "Priors for the long run", by Giannone, Lenza and Primiceri

Gianni Amisano (FRB) [1]

9th ECB Workshop on Forecasting Techniques: Forecast Uncertainty and Macroeconomic Indicators, Frankfurt, June 3-4 2016

[1] Views expressed here are not those of the FRS

1. What's the story
2. Some thoughts and questions on the methodology
3. Some thoughts and questions on the results
4. Conclusion

What's the story, I

- OLS estimated VAR models tend to attribute too much importance to deterministic trends in the explanation and therefore in the forecast path of the series
- This is a known problem: among others Sims (1996), (2000), Sims and Zha (1999)
- Known remedy: use priors that downplay importance of trend, e.g. sum of coefficient prior of Sims and Zha

What's the story

- But SZ prior does not work well if some of the variables being modelled are cointegrated
  - in a VECM representation, loadings on stationary linear combinations should be shrunk to zero more gently than those on non-stationary linear combinations
  - whereas SZ treats these two sets of loadings in the same way
- GLP (2016) propose a conjugate prior that does the job: the PLR prior

What's the story, III

- PLR generalises SZ, shrinking $A(1)$ to I_n but doing it with different intensities for stationary and non stationary combinations of the data
- a pre-sample is used to calibrate orior
- Computationally very convenient: all conjugate
- Applications with VAR models in different sizes (3, 5,7 variables)
- Shown better than SZ
- Also limitations of the proposed approach are shown

Importance of the prior on adjustment coefficients

- "In general, little attention has been given to the elicitation of informative priors on the adjustment coefficients, which is instead the main focus of our paper." (GLP, 2016, p. 15)
- Amisano and Serati (*Journal of Forecasting*, 1999): crucial how to set prior on adjustment coefficients

Deterministic trend and size of ρ

- Scalar AR(1) case with intercept
  $y_t = \left[c \times \frac{1-\rho^{t-1}}{1-\rho} + \rho^{t-1} \times y_1 \right] + \sum_{j=0}^{t-2} \rho^j \times \epsilon_{t-j}$
- The closer ρ to one, the simpler the deterministic component: when *rho* is one the trend is linear
- OLS estimates of ρ are downward biased
- priors pushing ρ towards one might do the trick
- But in multivariate framework SZ+Minnesota Prior are not sufficient in a potentially cointegrated model

Rank of Π

- Still a no cointegration prior, with Π being shrunk to zero and empirically, in finite samples, being full rank
- Would not a prior imposing rank reduction on Π be conceptually and maybe empirically preferable, in spite of being more complicated to implement?

Intercepts in cointegrating space

- Cointegration relationship with non zero mean, e.g. PPP relationship
- My guess is that these cases would require extra attention because mere size of $H_{.i} \times y_0$ will not be appropriate to measure how this relationship is tight in the pre-sample
- Hence decompose constant into two components, one in the cointegration space and the other out?

How to calibrate prior

- Rather than using relative size of $H_{.i} \times y_0$ to govern the shrinking
- Use pre-sample to compute serial correlation coefficients of linear relationships to calibrate shrinkage to zero
- A higher correlation coefficient in the pre-sample means slower convergence to equilibrium, hence requires stronger shrinkage

Using "wrong cointegration relationships

- EG: $c - y$: it seems very much at odds in US data
- What price do we pay in using a wrong relationship?
- Treat (some elements of) H as unknown and assign a prior?

Prior invariance with respect to rotations

- Interesting discussion in the paper
- An invariant prior is obtained by adding dummy observation to jointly shrink all non-stationary combinations together
- But this is invariant to rotations of the non-stationary combinations only
- How about rotations of stationary combinations?

Computed trends

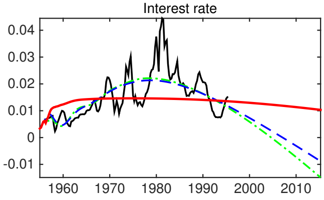

Figure: 2.1 inset

- Example of reaching ZLB catchy, but...
  - Recursively computed trends are bound to be very erratic
  - And most likely noisy: how "relevant" shall we consider the differences reported in the graph

In the three equation example

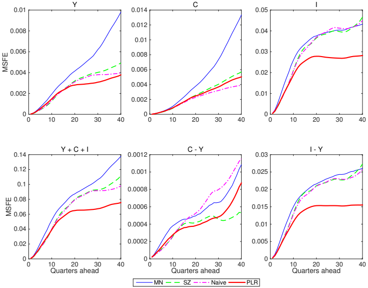

Figure: 5.1

- See Figure 5.1: for *y* and *c* differences between SZ and PLR quite negligible
- Is it consequence of a poor choice of cointegration relationship?

Size of the model and relative merits of priors

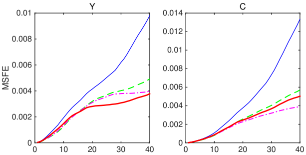

Figure: 5.1, inset

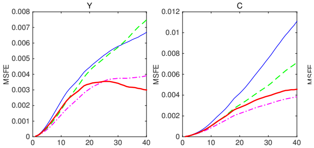

Figure: 5.3, inset

- Comparing Figures 5.1 and 5.3, moving from 3 to 5 variables and focussing on c and y again, it seems that the relative performance of SZ really deteriorates
- (Figure 5.5) This is further confirmed moving from 5 to 7 variables, with differences between SZ and PLR even more polarised
- Any intuition?

Conclusion

- Praise for the paper: very well written, simple idea and walking the reader through (most of) the relevant intuition
- LPR very simple to implement, but it requires some thinking. This is a very good thing
- It can be used in large information sets (I have some ideas for policy-related applications)
- Can be used as exploratory device and then use something more sophisticated
- I learnt a lot!