Discussion of "Forecasting low frequency macroeconomic events with high frequency data"

Lorenzo Trapani[1]

[1]University of Nottingham

4th November 2020

Modelling with Big Data Machine Learning: Measuring Economic Instability

Bank of England

The paper develops a methodology to predict (the probability of) an event S_t, defined as

$$S_t = \begin{array}{l} 1 \rightarrow \text{ there is a recession/ downside} \\ 0 \rightarrow \text{ there is no recession/downside} \end{array}.$$

This issue is interesting per se, obviously, and it can be viewed as a relatively rare event

- e.g. in the case of downside on financial markets, various papers set $y_t^* < -0.03$ (e.g. Farago and Tedongap, 2018, and Lettau, Maggiori and Weber, 2014), and a recent paper (Massacci, Sarno, Trapani, 2020) actually computes the threshold, which looks more like -0.06
- in fact, I would be quite interested in seeing more applications (not in this paper, I mean in general) to predict downsides in financial markets, maybe with UHF data.

Given that a probability is being modelled/predicted, a probit is used

$$P\left(y^*_{t+h} > 0\right) = \Phi\left(\beta_{0,h} + \beta_{1,h} z_t\right).$$

- This is not the main point of the paper, but I wonder whether other functional forms may be of help - if anything as robustness/sensitivity analysis

So, what is the main point of the paper? That whilst y_t^* is low-frequency, z_t is high frequency. If anything, this enables to have more info

$$P\left(y_{t+h}^* > 0\right) = \Phi\left(\beta_{0,h} + \sum_{k=1}^{K} \beta_{1,h,k} z_{t-(k-1)/m}\right);$$

the authors use, more parsimoniously, a weighted version of the lags $z_{t-(k-1)/m}$

$$P\left(y_{t+h}^* > 0\right) = \Phi\left(\beta_{0,h} + \beta_{1,h} \sum_{k=1}^{K} w_k z_{t-(k-1)/m}\right).$$

Dimensionality is always something to be wary of.

- Bayesian estimation is really great, because it allows for a lot of info on the uncertainty of predictions, but then computationally it can get tricky (am I right?);
- I am not an expert on Bayesian; I wonder how well the sampler works (does e.g. it mix well?)
- I wonder whether one cannot use something like a Bayesian compression algorithm, to get rid of some of the lags in $z_{t-(k-1)/m}$; or any shrinkage prior;
- The authors consider multiple predictors, i.e. z_t is N-dimensional. This reminds me of the large-VAR literature (speaking of Bayesian methods). Again, I would be curious to know what happens when N is large

- maybe here one could use Bayesian compression as I mentioned;
- possibly, although this is a nonlinear model, borrowing from the Bayesian VAR literature may help the obvious computational problems (e.g. Carriero et al., 2019; Tsionas, Izzeldin, Trapani, 2020 use copulas, which here could be more tricky - but see Kreuzer et al. 2019).
- There is a lot of insightful effort on measures of forecasting ability. I wonder whether combining forecasts, e.g. coming from different choices of predictors, may help?

Thank you!