Discussion on “Large Time-Varying Parameter VARs: A Non-Parametric Approach” by George Kapetanios, Massimiliano Marcellino and Fabrizio Venditti

Francesco Ravazzolo
Free University of Bozen/Bolzano

June 3, 2016
The paper introduces a nonparametric method for large VARs:
- It does not impose a specific form of time-variation.
- Estimators and asymptotic distributions are available in closed form.
- It allows for several types of shrinkage.
- Inference in terms of model selection criteria and pooling are provided.

The paper studies the properties of the new estimator in a simulation exercise.

Empirical applications:
- Point forecasting with 78 time-series.
- Response of industrial production indices to an unexpected increase in the price of oil.
In most of the paper, the proposed model assumes constant volatility.

Clark (2011) and Clark and Ravazzolo (2015) show that a (small size) constant parameter VAR with SV produces accurate forecasts.

Section 2.5 proposes a GLS estimator, but this is feasible only up to 20 variables.

This is not so bad, in particular considering that medium size VARs are often the most accurate.

Equation (31) requires the inversion of potentially large matrices \((nk \times nk)\).

- Block inversion.
- GPU.
The alternative model is the parametric model of Koop and Korobilis (2013).

- Allow for time-varying coefficients and time-varying variance/covariance matrix.
- Two further alternatives:
  - Time-varying coefficients but constant variance matrix (similar to the assumptions in the model presented in the paper).
  - Constant coefficients but time-varying volatility (Clark (2011), Clark and Ravazzolo (2015), Carriero, Clark and Marcellino (2016)).

Three DGPs:

- Time-varying coefficients follow a random walk with bounds on the first autoregressive parameter.
- Coefficients break only occasionally.
- Coefficients evolve as a sine function.
- All three cases assume stochastic volatility (and the nonparametric based model does not assume).
- Consider a specification with constant parameters and time-varying volatility. Interesting to learn how the nonparametric estimator functions in a similar case of misspecification.
Empirical application

- The paper focuses on point forecasting (RSPE).
- These models can provide larger gains in density forecasting.
Mean square prediction analysis

<table>
<thead>
<tr>
<th></th>
<th>fcst1</th>
<th>fcst2</th>
</tr>
</thead>
<tbody>
<tr>
<td>MSPE/Var(y)</td>
<td>1.000</td>
<td>1.013</td>
</tr>
<tr>
<td>VARIANCE</td>
<td>1.353</td>
<td>1.353</td>
</tr>
<tr>
<td>BIAS</td>
<td>0.001</td>
<td>0.003</td>
</tr>
</tbody>
</table>
Density forecast, model 2
Simulation example

DGP:

\[ y_t = N(\mu, \sigma), \; \mu = 0, \; \sigma = 1, \; t = 1, \ldots, T \]

Prediction models:

1. \( \hat{y}_{t,1} = N(\mu_1, \sigma_1), \; \mu_1 = 0, \; \sigma_1 = 1 \)
2. \( \hat{y}_{t,2} = N(\mu_2, \sigma_2), \; \mu_2 = 0, \; \sigma_2 = 2 \)
Structural analysis

- 28 variables VAR, 8 industrial production series.
- Response of the industrial production indices to an unexpected increase in the price of oil.
- Choleski decomposition (Edelstein and Kilian (2009)).
- No identification of oil supply and oil demand.
- Combination of identification strategies? Sign restriction for oil shocks only?
1 The $L_{mse}$ criterion considers a short window and discard values before it.
   - Why not a discounting factor?

2 Figure 1 shows that the optimized $\lambda$ hits the lower bound (1) in several occasions. Problems of convergence?