“Forecasting performance of three automated modeling techniques during the economic crisis 2007-2009”
by A. Kock and T. Teräsvirta:

a few comments





Éric Dubois
Directeur des Études et Synthèses Économiques

INSEE
INSTITUT NATIONAL DE LA STATISTIQUE ET DES ÉTUDES ÉCONOMIQUES

Plan of the discussion

1. A brief summary of the paper
2. A few (not very important) comments
3. A more important comment: non linearity versus structural break?
4. Another more important comment: to extend the functional set or the set of explanatory variables?

1. A brief summary of the paper

The paper deals with:

- the forecasting performance of three automated modeling techniques
- used to recover the best linear approximation of
- purely autoregressive
- non linear models
- with –almost– constant coefficients
- on 2 monthly series (IPI and unemployment)
- during the economic crisis 2007-2009

The main result:

The three examined methods have globally the same performance as the linear purely autoregressive model with –almost– constant coefficients

2. A few (not very) important remarks

1. What about real-time data?
2. Do we really want to discard all “insane” forecasts? What does the need to discard some “insane” forecasts tell us?
3. Is really direct non linear estimation less performing that the 3 automated methods applied to linear approximations?
4. Why choosing a very unusual 0.001 p-value for specifications test in *Autometrics*?
5. Please, could you be more precise on your testing device

2.1 What about real-time data?

1. The aim of the paper is to assess empirically the performance of forecasting methods with real data (presumably to help central bankers!)
2. Different methods should prove differently sensitive to revisions. The relative performance may therefore not be the same as in real time
3. So why did not you use real-time data?

2.2 The logic of the “insanity” filter

- In the paper, the authors sometimes “replace an unrealistic forecast with a more conventional and believable one” through an “insanity” filter
- More precisely, forecasts outside the historical range are replaced by RW or AR forecasts
- But, the crisis was a period outside historical range: we needed at this time “insane” forecast!
- If a method leads sometimes to truly insane forecasts, what guarantees that forecasts at other periods, although sane, are relevant?

2.3 Is really direct non linear estimation less performing?

- Direct non-linear estimation has been discarded because “it generally requires an iterative algorithm […] that may not behave well” (White, 2006)
- Why not test if it is truly the case that such direct estimation do not behave well?

2.4 Why the very unusual 0.001 p-value for specifications test in Autometrics?

- Why the choice of a very unusual 0.001 p-value for specifications test in Autometrics?
- Why not 0.01* (standard) or 0 (if there were estimation problems with 0.01)?

(*) with 5 specifications tests, (1-0.01)^5 # 0.95=1-0.05

2.5 Please could you be more precise on your testing device

- Not all details are available in the paper to recover the results (the vintage of the data, the precise estimation bounds, the presence of a constant, the software used, …)
- Readers would appreciate that your data and programs were made available on the web

3.1 A more important comment: non linearity versus structural break

- Is the crisis an event merely of bigger scale, but in line with previous experience? If this is the case, the advent of the crisis adds variance (a boon for applied econometricians!) and may help to uncover the true DGP, particularly when non linear
- Hence, the paper makes –implicitly– the hypothesis that the crisis is such an event in line with previous experience
- But this point merits discussion: does not the crisis mark a structural break, at least from the point of view of the dynamics of IPI and unemployment?

3.2 Non linearity versus structural break: a discussion

Many arguments in favor of the structural break hypothesis:

- What ever happened to the "great moderation"? Before the crisis, inflation expectations were well anchored and output was expected to fluctuate mildly around a smooth path; during the crisis, we have abruptly entered into a regime of "Knigthian" uncertainty

  ==> this has almost surely impacted the dynamics of production and unemployment (no reversion anymore to trend);
- Changes in credit distribution: before the crisis, (too) easy credit conditions; after the crisis, credit conditions much tighter

  ==> this can only have made firms more cautious and probably therefore has changed production dynamics (less smoothing for instance);
- Changes in capital allocation: before the crisis, wrong allocation of capital and of employment
- ==> after the crisis, the needed reallocations probably impacts the dynamics of unemployment (more persistence).

Forecasting performance of three automated modelling techniques during the economic crisis 2007": a few comments

3.3 Non linearity versus structural break: what do the data say?

Chow test in July 2008 for the model:

$$\dot{IPI}_t = a_0 + \sum_{i=1}^{i=6} a_i \dot{IPI}_{t-i}$$

With: $\dot{IPI}_t = \Delta(\log(IPI_t))$

And: IPI_t *the French Industrial Production Index*

Estimation period: Jan 1994-May 2010

Result : F(7,183) = 6.44

(p-value = 0.000000009)

What are the results of stability tests for the models estimated in the paper?

Forecasting performance of three automated modelling techniques during the economic crisis 2007": a few comments

3.4 Non linearity versus structural break: how to deal the problem?

If structural break is a prevalent problem, other methods may be needed (for instance, see Pesaran and Timmerman (2007), Pesaran et al. (2011) approach):

- Exponential smoothing;
- Averaging across estimation windows;
- Weighted least squares with weights robust to breaks…

Mixed, but overall encouraging, results on the data examined by Pesaran and al.: what results with IPI and unemployment data?

4. A more important comment: to extend the set of functions or of explanatory variables?

To improve the forecasting record for the IPI *vis-à-vis* linear AR models, should we:

- Extend the functional set beyond linearity? (the answer of the paper)

Or:

- Extend the set of explanatory variables?

At least in the French case, a –provisional – answer is: extend the set of explanatory variables, starting with business surveys.

4.1 The French case: the experimental design

- For i=1 to 31 estimation periods starting from 1995m7-2007m6 to 1998m1-2009m12,
  - Estimate the model: $\dot{IPI}_t = a_0 + \sum_{i=1}^{i=6} a_i \dot{IPI}_{t-i}$
  - Make recursive forecasts at horizons j=1 to 12
  - Store the corresponding forecasts into a (12 x 31) matrix
  - For each row of this matrix (truncated at 2009m12), calculate the corresponding RMSE
  - For j=1 to 12
    - Estimate the model:

      $\log(IPI_t / IPI_{t-j}) = a_0 + b_1 clim_{t-j+2} + b_2 \Delta clim_{t-j+2}$

      with: clim = manufacturing business climate
    - Make a direct forecast at horizon j
    - Store the corresponding forecast into cell (j,i) of another (12 x 31) matrix
  - For each row of this matrix (truncated at 2009m12), calculate the corresponding RMSE

4.2 The French case: results

relative RMSE of the model with business surveys*

horizon	relative RMSE
1	0.7367475
3	0.767772
6	0.7300975
12	0.9262500

* with respect to the pure AR one

4.3 The French case: extending the set of explanatory variables and functions?

- During the crisis, surveys have underestimated the fall of activity
- The case for non linear methods, such as the one exposed in the paper?
- A non linear model, derived from Cornec (2010), that is a non linearity of the following form:

$$\log(IPI_t / IPI_{t-j}) = a_0 + b_1 clim_{t-j+2} + b_2 \Delta clim_{t-j+2} + b_3 \Delta clim_{t-j+2} * |\Delta clim_{t-j+2}|$$

proves to be fruitful

==> It would be interesting to examine the performance of the more systematic techniques developed in the paper

4.4 The French case: the performance of Cornec non linear model

relative RMSE of the model with business surveys*

horizon	linear	non linear
1	0.7367475	0.7374579
3	0.767772	0.5782735
6	0.7300975	0.6164555
12	0.9262500	0.8836811

* with respect to the pure AR one

Remark: instability still remains (except at horizon 1)

Conclusion

- Forecasting during crisis times is a challenging task.
- Results for France suggest that business surveys should not be ignored to that end: but, is this the case for the other countries?
- Should we depart from the standard linear model?
  - Take into account non linearity?
  - Take into account structural breaks?

==> the matter of another paper?

Complements

Slides and programs to be made available (during the weekend) at:

http://dubois.ensae.net/biblio.html

All econometric results in this presentation use Grocer, the toolbox for Scilab (see:

http:// dubois.ensae.net/grocer.html)

References

- Cornec M. (2010): “Constructing a conditional GDP fan chart with an application to French business survey data”, 30th CIRET Conference, New York, October, https://www.ciret.org/conferences/newyork_2010/papers/upload/p_75-392908.pdf
- Pesaran M. H. , Pick A. and Pranovich M. (2011): “Optimal Forecasts in the Presence of Structural Breaks”, Working Paper, http://www.econ.cam.ac.uk/faculty/pesaran/wp11/PPP-30-Oct-2011.pdf
- Pesaran, M. Hashem, and Allan Timmermann (2007): ‘Selection of estimation window in the presence of breaks.’ *Journal of Econometrics* 137, 134–161.
- White (2006): “Approximate nonlinear forecasting methods, in G. Elliott, C. W. J. Granger and A. Timmermann (eds), *Handbook of Economic Forecasting*, Vol. 1, Elsevier, Amsterdam, pp. 459.512.

Thank you for your attention!

Contact
Éric Dubois
Tél. : (33) 1 41 17 60 82
E-mail: eric.dubois@insee.fr

Insee
18 bd Adolphe-Pinard
75675 Paris Cedex 14
www.insee.fr
Informations statistiques :
www.insee.fr / Contacter l'Insee
09 72 72 4000
(coût d'un appel local)
du lundi au vendredi de 9h00 à 17h00

INSEE