Discussion of "Nowcasting with Payments Data and Machine Learning"

Juan Antolin-Diaz[1]

"Modelling with Big Data and Machine Learning" Conference
Bank of England

November 26, 2018

[1]London Business School

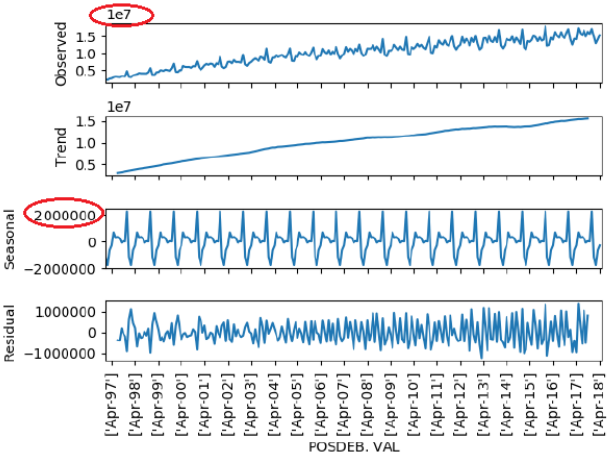
THIS PAPER

- Propose to use **payments data** to nowcast macroeconomic aggregates for Canada using **shrinkage** and **variable selection** techniques
- Daily payments data:
  - (Relatively) long sample: 1997-2017, covers two business cycles!
  - Very comprehensive: includes pre-authorised debits, small-value electronic payments, online payments
  - A number of disaggregated categories available
- Method:
  1. Aggregate payments data to monthly frequency
  2. Pre-filter payments and target series (RS and HHCE) to remove seasonality and trends
  3. Use Elastic Net (combination of Lasso and Ridge) to *estimate* the model (in-sample:1997-2013) select the best predictors using an expanding *out-of-sample* window (2014-2017)

General Issues

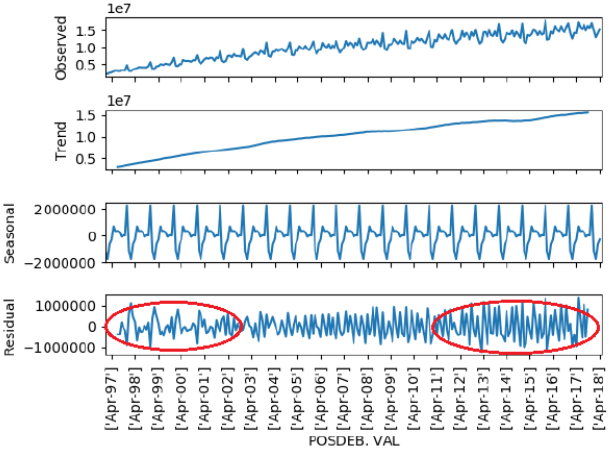
- Curious why the target series are retail sales and household consumption. Maybe interesting in their own right but decision makers are typically interested in broader measures of economic activity such as GDP.
- With disaggregated payments data one cannot avoid the fear that technological progress / preference shifts may render the relationship between targets and predictors highly unstable. Why not just use the aggregate series?

Data Pre-Processing: Concerns



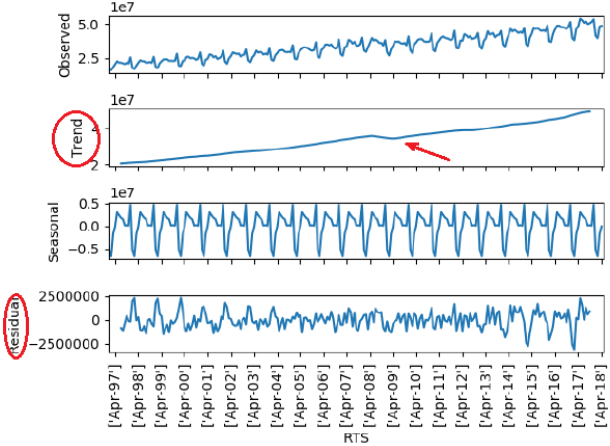
- With economic time series that grow exponentially, usually a good idea to work with natural logarithms.

Data Pre-Processing: Concerns



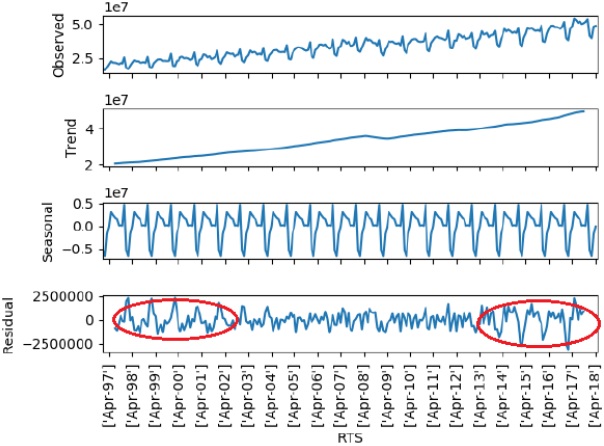
- Evidence of heteroskedasticity and changing seasonality.

Data Pre-Processing: Concerns



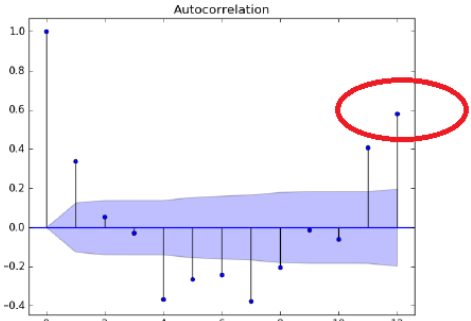
- Removing the trend (includes business cycle frequency) throws away the one component we *are* interested about.

Data Pre-Processing: Concerns

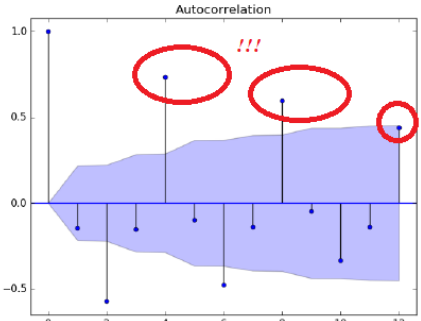


- Clear evidence of residual seasonality. Using tried-and-tested packages for economic time series (X12/13, Tramo-Seats) probably a good idea.

Data Pre-Processing: Concerns



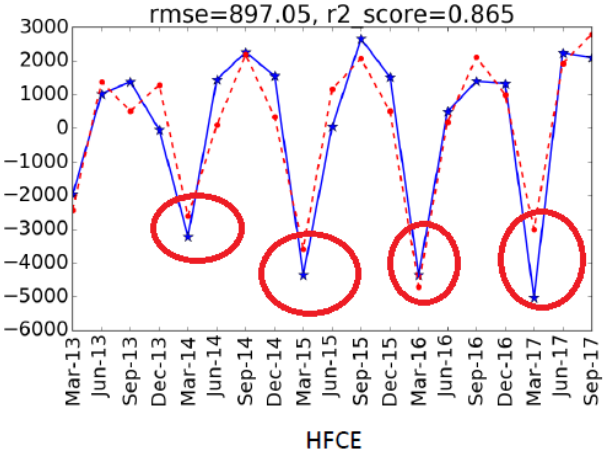
RTS
(Monthly)



HFCE
(Quarterly)

- Clear evidence of residual seasonality. Using tried-and-tested packages for economic time series (X12/13, Tramo-Seats) probably a good idea.

Results: Concerns



- Spurious over-fitting of residual seasonality?

An alternative

Compare to state of the art specification using timely data

Antolin-Diaz, Drechsel and Petrella (2017, 2018): Bayesian Dynamic Factor Model with TVM, SV and automatic outlier detection:

$$\Delta(\mathbf{y}_t - \mathbf{o}_t) = \mathbf{c}_t + \mathbf{\Lambda}(\mathbf{L})\mathbf{f}_t + \mathbf{u}_t, \quad (1)$$

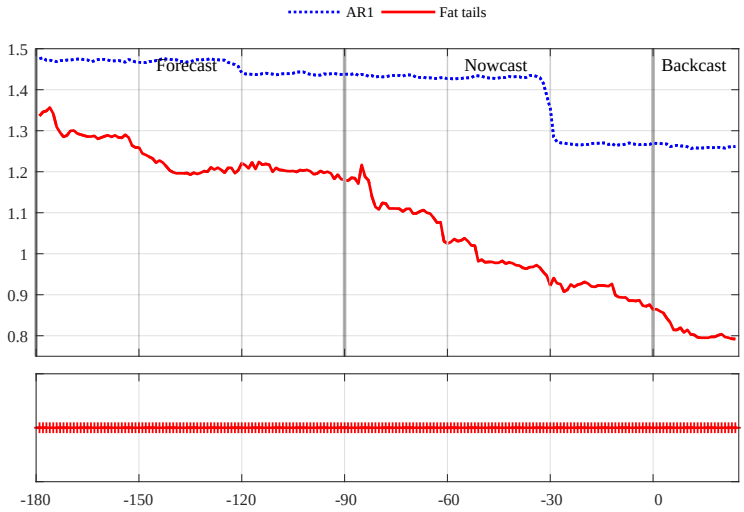
$$(I - \Phi(L))\mathbf{f}_t = \sigma_{\varepsilon_t}\boldsymbol{\varepsilon}_t, \quad (2)$$

$$(1 - \rho_i(L))u_{i,t} = \sigma_{\eta_{i,t}}\eta_{i,t}, \quad i = 1, \ldots, n \quad (3)$$

Evaluated over a 16-year out-of-sample window using fully real time (unrevised) data.

RESULTS: DENSITY FORECASTING (CANADA)

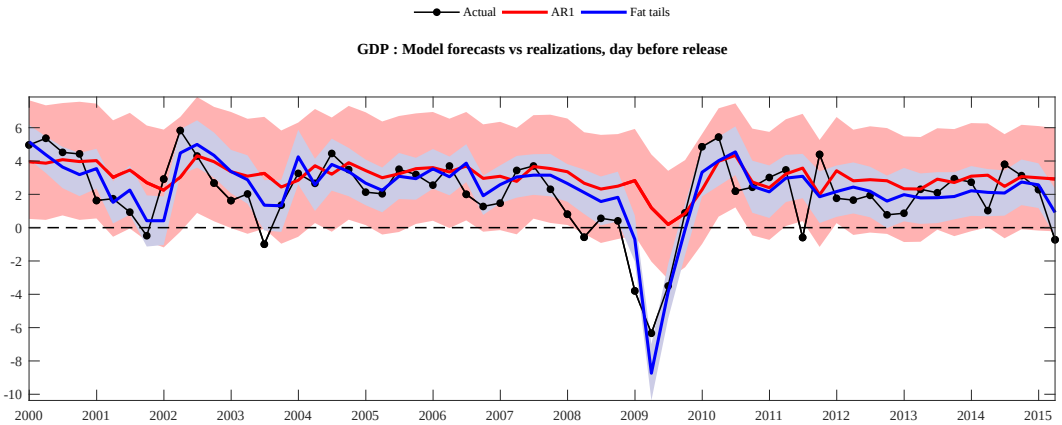
GDP: CRPS ACROSS HORIZONS



- Nowcasting is about the impact of the incoming data flow on our (probabilistic)

EVALUATION RESULTS

FORECASTS VS. ACTUAL OVER TIME (CANADA)



- The key question is whether payments data improves (and at what point in the release calendar) in a model that already includes the most timely macro data (i.e. business surveys).

FINAL COMMENTS ON METHODOLOGY

- Many macro data sets are actually quite *small* both in the N and T dimension. Careful analysis and understanding of the properties of the data is *feasible* and usually *pays off*.
- We know that ML methods work well on very large cross-sectional data sets. More research is needed on their properties for *small* panels of *highly correlated* time series.
- Giannone, Lenza, Primiceri (2018): little evidence of models with few predictors in Macro. "A clearer pattern of sparsity can only emerge when the researcher has a strong a-priori bias in favor of predictive models with a small number of regressors".

Thanks

References

- Antolin-Diaz, J., Drechsel. T., and Petrella, I., (2017) "Tracking the slowdown in Long-Run GDP growth," The Review of Economics and Statistics, MIT Press, vol. 99(2), pages 343-356, May.
- Antolin-Diaz, J., Drechsel. T., and Petrella, I., (2018) "Advances in Nowcasting Economic Activity," work in progress.
- Giannone, D., Lenza, M., and Primiceri, G., (2018) "Economic Predictions with Big Data: The Illusion of Sparsity", CEPR working paper DP12256.