ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

LIII Reunión Anual

Noviembre de 2018

ISSN 1852-0022
ISBN 978-987-28590-6-0

Nowcasting Activity in Argentina using Dynamic Factor Models

Blanco Emilio Fernando
D'amato Laura
Dogliolo Fiorella
Garegnani María Lorena

Nowcasting Activity in Argentina using Dynamic Factor Models*

Emilio Blanco
BCRA and UBA

Laura D'Amato
BCRA and UBA

Fiorella Dogliolo
BCRA and UNLP

Lorena Garegnani
BCRA and UNLP

August 2018

Abstract

Having a correct assessment of current business cycle conditions is one of the major challenges for monetary policy conduct. Given that GDP figures are available with a significant delay central banks are increasingly using Nowcasting as a useful tool for having an immediate perception of economic conditions. We develop a GDP growth Nowcasting exercise using a broad and restricted set of indicators to construct different models including dynamic factor models as well as a FAVAR. We compare each individual model with alternative combinations in a pseuro-real time out-of-sample exercise and find an improvement in predictive performance using the Giacomini and White (2004) test. Finally we introduce a DFM state-space approach, being able to measure the impact of data releases or news on sequential forecast.

Keywords: Nowcasting, state-space, DFM
JEL classification: C22, C53, E37

*We thank Domenico Giannone, Peter Zadrozny and Marcos Dal Bianco for their insightful comments to earlier versions of this paper. We especially thank Chen Baolin for her invitation to the SEM's 2017 Conference at MIT Cambridge, Massachusetts. We also thank seminar participants at CIRET Conference 2016, SEM's 2017 Conference and 2017 AAEP Annual Meeting. The opinions expressed in this work are those of the authors, and do not necessarily reflect those of the Central Bank of Argentina authorities. Email: emilio.blanco@bcra.gob.ar; ldamato@bcra.gob.ar; fiorella.dogliolo@bcra.gob.ar; lgaregnani@bcra.gob.ar

1 Introduction

Although the assessment of current economic conditions is a crucial ingredient of decision making in central banks and other areas of the government, this process has to be conducted in real time based on incomplete information, mainly because Gross Domestic Product (GDP) -the main source of information on economic activity-is released on a quarterly basis and with an important lag. At the same time, a large number of business cycle indicators are available at higher frequencies as monthly or even daily.

Nowcasting -defined as the prediction of the present, the very near future and the very recent past (Giannone et al., 2008, Banbura et al., 2013) - has proved to be a useful tool to overcome this problem. As a result, its use by central banks and other government institutions has been growing rapidly over the recent years. A contraction for *now* and *forecasting*, Nowcasting is a technique mostly applied in meteorology which has been recently introduced in economics. Its basic principle is the exploitation of the valuable information content embodied in a large number of business cycle indicators that are available at high frequencies -daily or monthly- to produce early estimates of a target variable published at a lower-quarterly-frequency. This early estimations can be sequentially updated, when new information becomes available.

The most simple and earlier version of Nowcasting are Bridge equations, i.e. a combination of simple bivariate autoregressive models (Kitchen and Monaco, 2003; Drechsel and Maurin, 2008). Recently new statistical approaches where developed to overcome problems inherent to Nowcasting. Dynamic factor models (DFM) (Stock and Watson, 2002, 2006), implemented through the estimation of principal components or a state space representations (Evans, 2005; Giannone et al., 2008; Arouba, et al., 2009) address the *high dimension problem*[1] extracting common factors to large sets of indicators. Mixed Data Sampling (MIDAS) equations (Ghysels et al., 2004) and state space representations of DFMs provide solutions to the *mixed-frequency problem*.[2] All of them have proved to be effective in anticipating short-term developments. They also seem to outperform univariate statistical models in terms of the predictive performance, particularly in volatile environments (Bell et al., 2014).

Giannone et al. (2008) highlight as the main advantages of *Nowcasting*: **(i)** The use of a large number of data series, from different sources and frequencies; **(ii)** the updating of estimates as new information becomes available (in accordance with the real-time calendar of data releases) and **(iii)** the fact that it "bridges" monthly data releases with quarterly GDP.

Real-time short term forecasting involves typically two type of business cycle variables: **(i)** *Hard indicators* of economic activity -such as industrial production and its components, housing indicators, energy consumption and production and financial and monetary time series as money aggregates, interest rates and **(ii)** *Soft indicators* mostly coming from surveys which mainly reflect agents' perceptions about economic conditions, as consumers confidence indexes.

In this paper we consider a broad set of different *Nowcasting* models and conduct a pseudo-real-time one quarter ahead forecasting exercise to predict quarterly GDP growth figures in Argentina over the period 2006:Q1-2017:Q1. The task is particularly challenging because the economy was subject to several shocks over this years, including two sharp currency depreciations in January 2014 and December 2015. Since then the economy has been going through major structural reforms including the adoption of a floating exchange regime, the removal of exchange rate controls and the adoption of an inflation targeting scheme to conduct monetary policy.

[1] Also known as the curse of dimensionality or parameter proliferation, where there is a large number of parameters relative to the number of observations

[2] Time series involved are sampled at different frequencies and published at dissimilar lags (unbalanced data sets).

Our main objective is to compare the relative predictive ability of the following *Nowcasting* models: **(i)** a Factor Model, **(ii)** a Factor Model in which we split the dataset in three factors (*hard*, *soft* and *price* indicators) and **(iii)** a Factor augmented VAR (FAVAR). We also evaluate the potential gain of using a large group vs a reduced group of business cycle indicators considering two data groups: one that includes 112 indicators and a more restricted subset of 22 indicators (those that have the highest contemporaneous correlation with GDP growth).[3] We conduct a Giacomini and White test (2004) to assess the statistical significance of potential differences in forecasting performance. The finding that none of them outperforms the rest of models led us to investigate whether a forecast combination could perform better than individual models.

Stock and Watson (2016) show that both parameters and factors of a DFM can be estimated using nonparametric methods (such as principal components analysis) or by parametric state-space methods. We first follow a PCA approach, averaging monthly data to convert it to a quarterly frequency and then obtain the factors to forecast GDP at its original frequency. Additionaly we present a state-space version of the Dynamic Factor Model **(i)** in the spirit of Banbura, Giannone and Reichlin (2010). An improvement of applying this methodology is that averaging to a lower frequency is no longer necessary as monthly GDP figures can be estimated as an unobservable variable through the kalman filter. Although preliminary[4] results are promising.

The paper is organized as follows. The data set and our first empirical approach are presented in section 2, along with the results obtained from the *Nowcast* exercise. Section 3 complements this first exercise through a forecast combination approach.The fourth section evaluates the relative predictive ability of the *Nowcasting* models using the Giacomini and White (2004) test. Next, we introduce a state-space Dynamic Factor model, describe main advantages and present preliminary results for this methodology. Finally, section 6 concludes.

2 Our First Nowcast Exercise

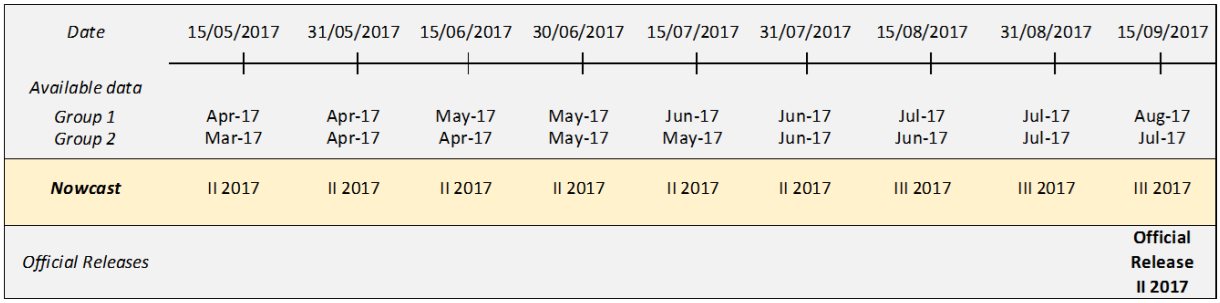
Our exercise consists on producing early predictions of GDP growth based on the sample period 2006:Q1 - 2017:Q1. In Argentina the official GDP figures are released around 10 weeks after the end of the quarter. The initial data set comprises 112 business cycle indicators, including *hard* and *soft* business cycle time series, ranging from financial indicators to tax collection data, disaggregated data on industrial production, consumer confidence surveys and car sales. The variables comprised in the data set are described in Annex 1. The series were seasonally adjusted (when needed) using X-13 ARIMA-SEATS, detrended or differentiated to make them stationary and finally log transformed. Using an estimation sample that comprises the period 2006:Q1-2010:Q4, we perform rolling pseudo-real-time one quarter ahead *Nowcast* exercise of GDP growth over the period 2011:Q1-2017:Q1 with a window size of *20 quarters*, using the methodologies described below for two sets of series: A broad one, composed by 112 business cycle indicators and a subset of 22 series that exhibit the highest contemporaneous correlation with GDP growth within the unrestricted set (see Annex 1).

According to the timing of publication we split the final set of indicators in two groups: those series that are available less than 10 days after the end of each month (Group 1), and series that are published with a delay raging form 10 to 30 days (Group 2). Following this grouping of the series, the Nowcast can be sequentially updated as described in Figure 1.

[3] In this regard, Boivin and Ng (2006) find that a factor model forecast derived from 40 time series performs better than a 147 series factor model. Similar results can be found in Caggiano, Kapetanios and Labhard (2009).

[4] We only considered the restricted set of indicators and work with a short sample

Figure 1: Sequential updating example



As reported by the aforementioned updating scheme, we can obtain 6 early estimations of the GDP growth within each quarter.

2.1 The methodological approach: Factor Models through PCA

Nowcast can be conducted through the estimation of common factors from a large set of monthly data and subsequently using them as regressors for GDP -as proposed by Giannone, Reichlin and Small (2005).The idea behind this approach is that the variables in the set of interest are driven by few unobservable factors.

More concretely, the covariance between a large number of n economic time series with their leads and lags can be represented by a reduced number of unobserved q factors, with $n > q$. Disturbances in such factors could in this context represent shocks to aggregate supply or demand.

Therefore, the vector of n observable variables in the cycle can be explained by the distributed lags of q common factors plus n idiosyncratic disturbances which could eventually be serially correlated, as well as being correlated among i.

A vector X_t of n stationary monthly business cycle indicators $x_t = (x_{1t}, ..., x_{nt})'$, with $t = 1, ...T$ can be explained by the distributed lags of q common latent factors plus n idiosyncratic disturbances which could eventually be serially correlated. The dynamic factor model (DFM) is therefore:

$$X_t = \lambda(L) f_t + e_t \tag{1}$$

$$f_t = \Psi(L) f_{t-1} + \eta_t \tag{2}$$

Where f_t is a vector $q \times 1$ of unobserved (latent) factors, the lag polynomial matrices $\lambda(L)$ and $\Psi(L)$ are $N \times q$ and $q \times q$, η is a $q \times 1$ vector of (serially uncorrelated) innovations to the factors. The i row of $\lambda(L)$ is called the *dynamic factor loadings* for the ith series, X_{it}, and the e_t are the idiosyncratic disturbances that are assumed to be uncorrelated with the factors in all leads and lags[5], that is to say $E(f_t e_{it}) = 0 \; \forall \; i, s$.

Given a target variable y_t (in our case GDP growth), the objective would be to estimate $E(y_t|X_t)$

$$y_t = \beta_t X_t + \gamma_t y_{t-1} + \varepsilon_t \tag{3}$$

$$y_t = \beta(L)' f_t + \gamma(L)' y_{t-1} + \varepsilon_t \tag{4}$$

[5]See Stock and Watson (2016) for cases where e_i is serially correlated

If the lag polynomials $\lambda(L)$ in (1) and $\beta(L)$ in(4) are of finite order p [6], Stock and Watson (2002a) show that the factors f can be estimated by **principal components**.

If we define quarterly GDP as the average of monthly latent observations $y_t^Q = (y_t + y_{t-1} + y_{t-2})$ and we obtain quarterly factors f_t^Q from these observations, we can use the following bridge equation to obtain early estimates of GDP:

$$\widehat{y_t}^Q = \beta(L)' f_t^Q \tag{5}$$

In addition to estimate models using a single equation approach as in (4), we also calculate a VAR on GDP growth and the factors (FAVAR). [7]

2.2 Results

To estimate the factors for both the restricted and unrestricted sets of indicators we proceeded in the following way. We use the two sets of indicators to calculate the factors using the principal component methodology. Then we determine the number of factors to be used to estimate the models using *scree plots*[8]. Based on this we use the factors of all the models detailed above, from (M1) to (M6) the restricted and the unrestricted sets of indicators.

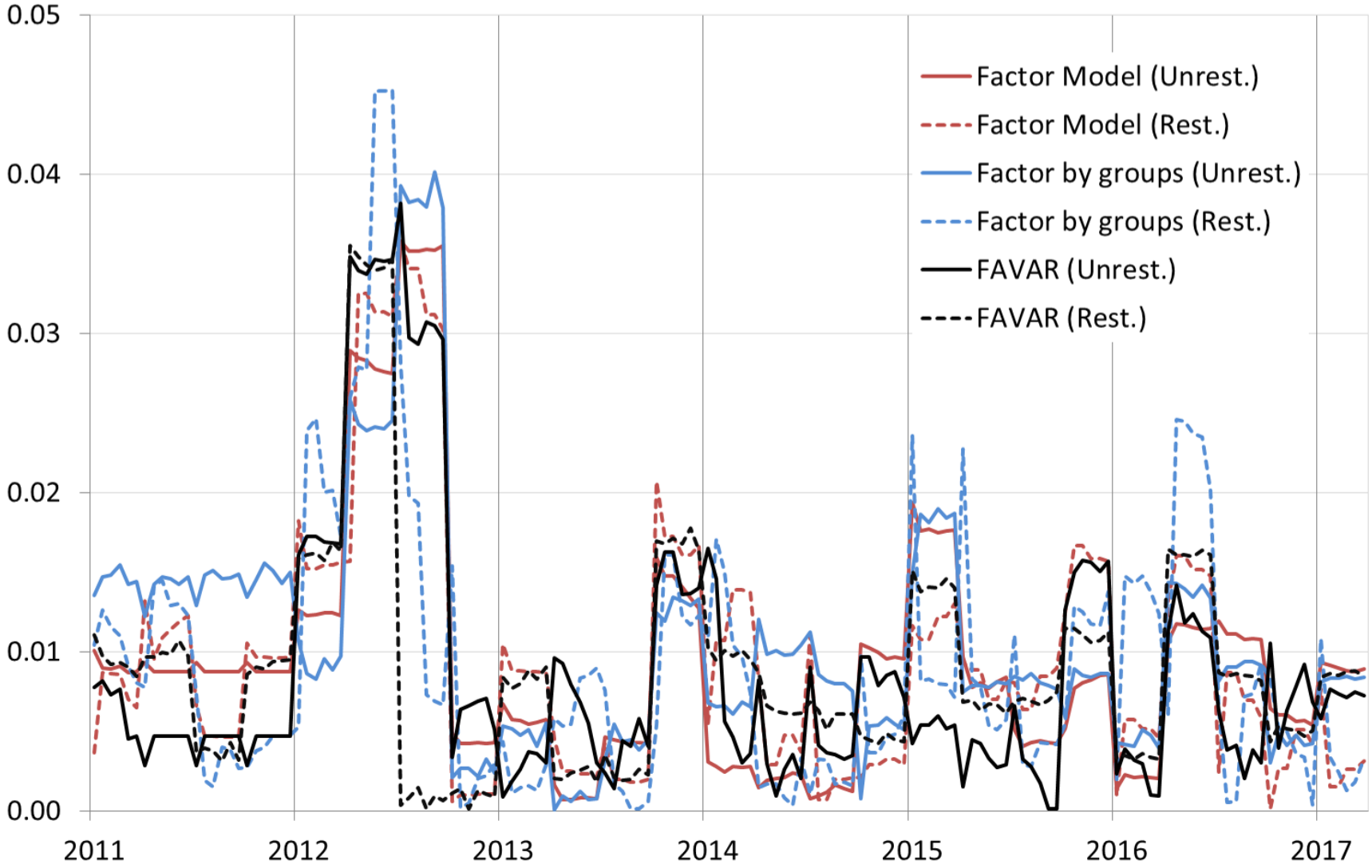
Based on these estimations we conduct a rolling window Nowcasting exercise for each of the models over the period 2011:Q1-2017:Q1. Figure 2 presents for each quarter of the predictive sample the correspondent loss function for each model, measured by the Root Mean Square Error (RMSE). At first sight it seems that none of the models outperforms the rest.

[6] Expressing the DFM in a static (or stacked) form

[7] Estimation results are available upon request.

[8] Developed by R B. Cattel in "The scree test for the number of factors", Multivariate Behav. Res. 1:245-76, 1966.University of Illinois, Urbana-Champaign, ILl.

Figure 2: Nowcast performance (RMSE)



To provide a better insight, Figure 3 shows the frequency at which each model is ranked as first in terms of its forecast accuracy (measured by the RMSE).[9] We calculate these frequencies for the complete sample and then we split it into the pre and post structural break in December 2015. The FAVAR models outperforms the rest of the models for the complete predictive sample, while in the last period (although the sample is quite short) it seems that other models, as the restricted and the unrestricted factor models, have rather the same predictive ability. To verify if the observed difference in predictive ability are statistically significant we conduct the Giacomini White test as described in Section 5.

Figure 3: Frequency at which models are ranked as first in terms of predictive accuracy (lowest RMSE)

% of periods with lowest RMSE	*Full Sample* 2011q1-2017q1 (150 obs)	1^{st} *Sample* 2011q1-2015q4 (120 obs)	2^{nd} *Sample* 2016q1-2017q1 (30 obs)
M1: Factor Model (Rest.)	13%	9%	**27%**
M2: Factor Model (Unrest.)	17%	17%	20%
M3: Factor by groups (Rest.)	25%	24%	**27%**
M4: Factor by groups (Unrest.)	11%	13%	0%
M5: FAVAR (Rest.)	7%	8%	0%
M6: FAVAR (Unrest.)	**28%**	**28%**	**27%**

3 Pooling of Nowcasts

Since Bates and Granger (1969), the forecasting literature has emphasized that a combination of different forecasts might result in a better performance in comparison with each individual model. This technique is particularly useful in the presence of structural breaks. The pooling or combination of forecasts implies combining two or more forecasts derived from models that use different predictors to produce a forecast. The basic idea is as follows:

Let $\left\{ \sum_{i,t+h}^{h} Y; i = 1...n \right\}$ be a panel of n forecasts. The combined forecast or forecasting pool will be given by the linear combination:

$$Y_{t+h|t}^{h} = \omega_0 + \sum_{i=1}^{n} \omega_{it} Y_{i,t+h}^{h} \tag{6}$$

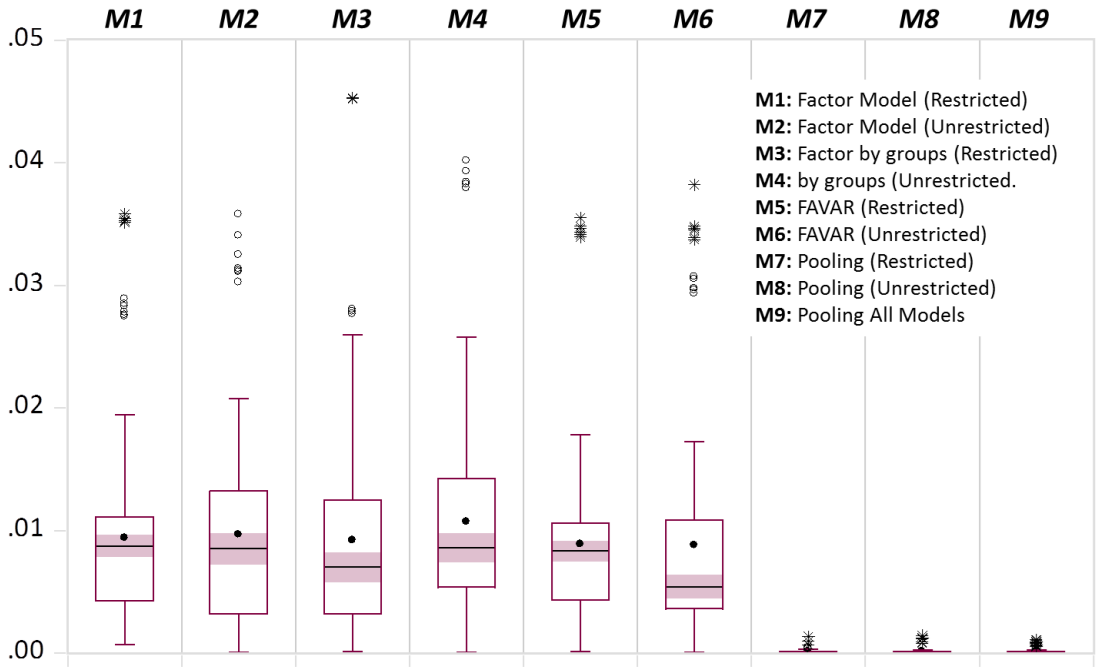
where ω_{it} is the weight of the i^{th} forecast in period t.[10]
Given this general setting, we conduct a pooling exercise of the different forecasting models using equal weights:

- (M7) Combining only Restricted models
- (M8) Combining only Unrestricted models
- (M9) Combining all models

Figure 4 depicts the box-plots of the RMSE distribution for all Factor Models and the three forecast combinations described above. It can be seen that any of the forecast combinations seems to outperform any of the individual models. In the next section we evaluate the differences in predictive ability of all of the nowcasting exercises reported in Figure 4.

[9] See Annex 2 for a relative position histogram of each of the models for the full sample.
[10] See D'Amato et al. (2009) for a brief discussion on different weighting schemes.

Figure 4: Box plots of RMSE



4 Testing for equal predictive ability

To test if the differences in predictive accuracy found in the previous section are statistically significant we use the Giacomini and White (2004) test. The Giacomini and White approach differs from that followed by previous tests, as those proposed by Dieblod and Mariano (1995) and West (2003) in what it is based on conditional rather than unconditional expectations. In this regard, the Giacomini and White approach focuses on finding the best forecast method for the following relevant future. Their methodology is relevant for forecasters who are interested in finding methodologies that improve predictive ability of forecast, rather than testing the validity of a theoretical model.[11]

The test has many advantages: **(**i**)** it captures the effect of estimation uncertainty on relative forecast performance, **(**ii**)** it is useful for forecasts based on both nested and non nested models, **(**iii**)** it allows the forecasts to be produced by general estimation methods, and **(**iv**)** it is quite easy to be computed. Following a two-step decision rule that uses current information, it allows to select the best forecast for the future date of interest.

The testing methodology of Giacomini and White consists on evaluating forecast by conducting an exercise using rolling windows. That is, using the R sample observations available at time t, estimates of y_t are produced and used to generate forecast τ step ahead. The test assumes that there are two methods, f_{Rt} and g_{Rt} to generate forecasts of y_t using the available set of information $\mathcal{F}_t$. Models used are supposed to be parametric.

$$\begin{aligned} f_{Rt} &= f_{Rt}(\widehat{\gamma}_{R,t}) \\ g_{Rt} &= g_{Rt}(\widehat{\theta}_{R,t}) \end{aligned}$$

A total of P_n forecasts which satisfy $R + (P_n - 1) + \tau = T + 1$ are generated. The forecasts are

[11] See Pincheira (2006) for a nice description and application of the test.

evaluated using a loss function $L_{t+\tau}(y_{t+\tau}, f_{R,t})$, that depends on both, the realization of the data and the forecasts. The hypothesis to be tested is:

$$\begin{aligned} H_0 \quad &: \quad E\left[h_t\left(L_{t+\tau}(y_{t+\tau}, f_{R,t}) - L_{t+\tau}(y_{t+\tau}, g_{R,t})\right) \mid \mathcal{F}_t\right] = 0 \\ &\quad \text{or alternatively} \\ H_0 \quad &: \quad E\left[h_t \Delta L_{t+\tau} \mid \mathcal{F}_t\right] = 0 \quad \forall\, t \geqslant 0 \end{aligned}$$

for all $\mathcal{F}_t$ -measurable function h_t.

In practice, the test consists on regressing the differences in the loss functions on a constant and evaluating its significance using the t statistic for the null of a 0 coefficient, in the case of $\tau = 1$. When τ is greater than one, standard errors are calculated using the Newey-West covariances estimator, that allows for heteroskedasticity and autocorrelation.

The results of applying the Giacomini and White procedure to evaluate the forecasting performance of the models are shown in Figure 5. They corroborate the intuitions provided by the descriptive analysis in the previous section: While none of the individual models outperforms the rest of them, all the forecast combinations perform much better than individual models. However, one must notice that there is no clear prevalence of any particular combination over the others.

Figure 5: Results of the Giacomini and White test

Difference in RMSE Model in row - Model in Column

G&W test	**M1:** Factor Model (Rest.)	**M2:** Factor Model (Unrest.)	**M3:** Factor by groups (Rest.)	**M4:** Factor by groups (Unrest.)	**M5:** FAVAR (Rest.)	**M6:** FAVAR (Unrest.)	**M7:** Pooling (Rest.)	**M8:** Pooling (Urest.)	**M9:** Pooling All
M1: Factor Model (Rest.)		0.000	0.000	-0.001	0.001	0.001	0.0095	0.0095	0.0095
M2: Factor Model (Unrest.)	-0.0003		0.000	-0.001	0.000	0.001	0.0093	0.0092	0.0092
M3: Factor by groups (Rest.)	-0.0005	0.000		-0.002	0.000	0.002	0.0091	0.0090	0.0091
M4: Factor by groups (Unrest.)	0.0010	0.001	0.002		0.002	0.000	0.0106	0.0105	0.0106
M5: FAVAR (Rest.)	-0.0009	0.000	0.000	-0.002		0.000	0.0088	0.0087	0.0088
M6: FAVAR (Unrest.)	-0.0008	-0.001	0.000	-0.002	0.000		0.0086	0.0086	0.0086
M7: Pooling (Rest.)	-0.0095	-0.0093	-0.0091	-0.0106	-0.0088	-0.0086		0.0000	0.0000
M8: Pooling (Urest.)	-0.0095	-0.0092	-0.0090	-0.0105	-0.0087	-0.0086	0.0000		0.0000
M9: Pooling All	-0.0095	-0.0092	-0.0091	-0.0106	-0.0088	-0.0086	0.0000	0.0000	

Significant at 5%

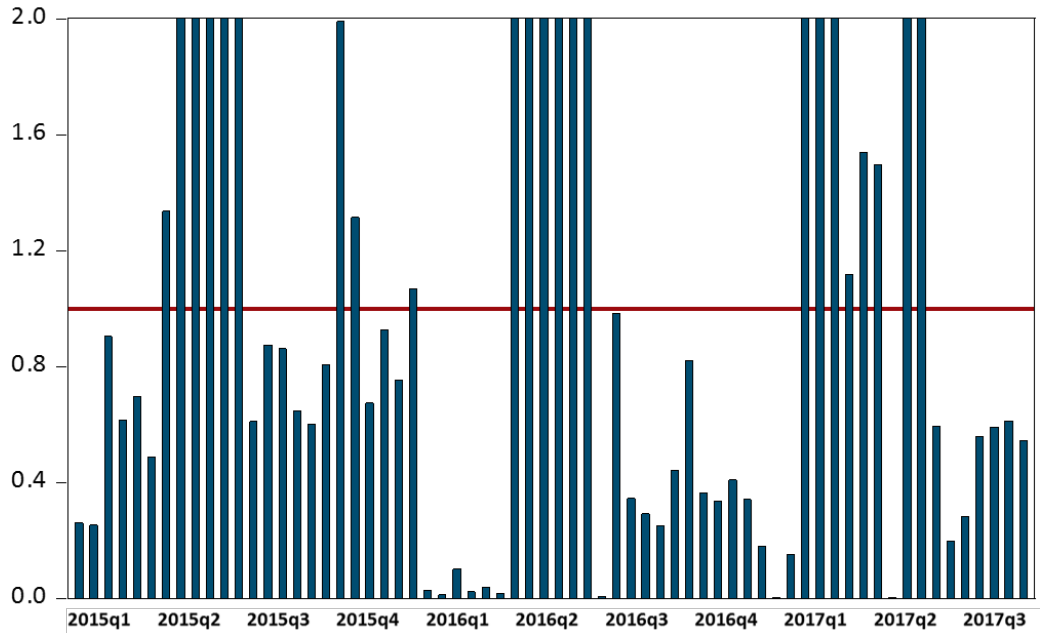
5 Second Nowcasting exercice: a State-Space Respresentation of DFM

Following Banbura, Giannone and Reichlin (2010) and Stock and Watson (2016), we might also present (1) and (2) in a State-Space form in order to compute the likelihood via the Kalman Filter. One common working assumption to do it is to suppose that the errors idiosyncratic component e_t follows a independent univariate autoregression ($e_{it} = \delta_i(L)e_{it-1} + \nu_{it}$, with $\nu_{it} \sim N(0, \sigma^2_{\nu i})$).

Under this setting, the first equation of the DFM (1) is the measurement equation, relating observed data (in our case the vector X of 22 Business cycle indicators) and the state (latent) variables (the underlying factor or common component). On the other hand, equation (2) is the transition equation describing the dynamics of the state variables. The final step is to consider Mariano and Murasawa (2003) suggestion to incorporate the low frequency (quarterly) target variable into this framework as a partially observed monthly indicator. Defining θ as a vector that incorporates all the parameters of the model (the λ factors loadings, ε_t and σ_t). Once the joint model is set up in State-Space form, like Banbura *et al* (2010) we estimate the parameters θ of the state space form by the Expectation Maximisation (EM) algorithm.[12]

Figure 5 compares the RMSE of this method with a PCA factor approach. A bar below the red line indicates that on that particular estimation the State-Space DFM has a lower forecast error: this happens in 65% of the cases. When evaluating if both methods are different in terms of foracast accuracy the Giacomini and White test is not able to reject the null of equal predicitve preformace. It should be noted, however, that given the presence of a structural break in December 2015, this result should require further research to determine to what extent the relative performance of the two model could have changed over time (Giacomini and Rossi, 2010), using for example the the Giacomini and Rossi Fluctuation test.

Figure 5:
RMSE ratio (State-Space DFM/Nowcast)

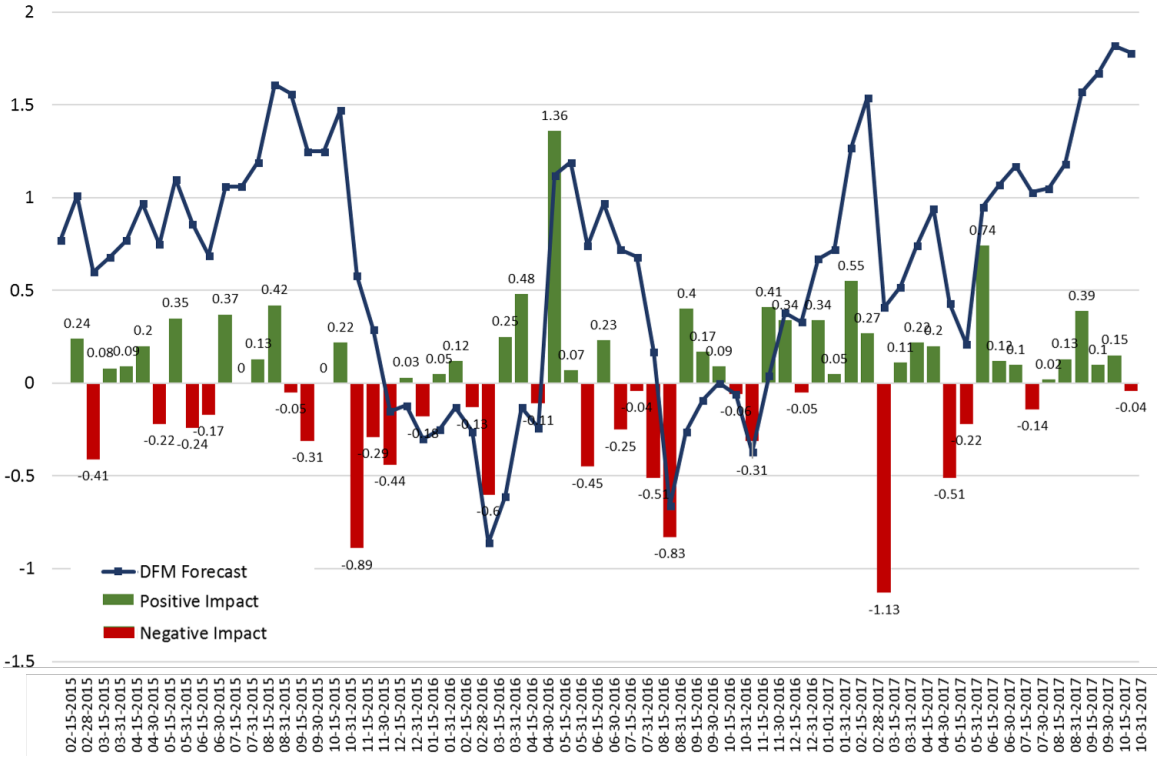


Regardless of this preliminary result, a major advantage posed by this approach is working with mixed frequency data: unlike our first exercise we can work in a high (monthly) frequency. According

[12] See Banbura et al (2010) for further details.

to Stock and Watson, the state space framework is flexible enough to solve the ragged-end problem allowing the measurement equation to cope with missing observations at different dates t.

Figure 6: News impact



The solid blue line is the quarterly GDP change forecast, while the vertical bars indicate whether the impact of new data was positive (green) or negative (red). [13] In terms of communication this a breakthrough from previous methodologies.

The joint estimation of the latent factors and the low frequency variable, allows for an estimation of a GDP at a monthly frequency. [14]

6 Concluding Remarks

One of the main concerns of monetary policy should be taking decisions based on *real-time* assessment of current and future business cycle conditions. Nevertheless in practice, Gross Domestic Product (GDP), the main source of information on economic activity, is released in Argentina on a quarterly basis and with a 10 week lag.

Nowcasting might be useful to overcome this problem. However, a mayor dilemma faced when working in a rich-data environment is that data are not all sampled at the same frequency. In recent years, the forecasting literature has developed a series of solutions to deal with this *mixed-frequency problem*.

We conduct a nowcasting exercise for GDP growth in Argentina over the period 2006:Q1 -2017:Q1 for a bunch of different nowcasting models including a Factor model and a FAVAR for both a large as well as a more restricted set of business cycle indicators. The exercise is quite challenging because

[13] One might also compute the contribution of different Categories of information. We are currently working on this regard.

[14] Some evidence on this regard is presented in Annex 3

the Argentine economy is currently experiencing a structural break. The results indicate that there is no statistically significant prevalence of a model over the others in terms of predictive ability while there seem to be some gains of combining them to produce nowcast.

Although preliminary, introducing a State-Space approach can help us to better communicate results to both *policy makers* and the public, as well as solve some of the common issues faced when working with mixed frequency data. In our particular case it also improves the treatment of mixed frequency relative to the PCA Factor Model (averaging to transform monthly to quarterly data is no longer necessary). Further research agenda consist in including this type of model into a combination exercise, improving the series selection process and apply the Giacomini and Rossi methodology to test differences in forecasting ability in unstable environments.

References

[1] Andreou E., Ghysels, E. and A. Kourtellos (2012). "Forecasting with mixed-frequency data". Chapter prepared for *Oxford Handbook on Economic Forecasting* edited by Michael P. Clements and David F. Hendry.

[2] Angelini, E., Camba-Méndez, G., Gianonne, D., Rünsler, G. and L. Reichlin (2008). "Short-term forecast of Euro Area GDP". European Central Bank Working Paper No. 949.

[3] Armesto M., Engemann, K. and M. Owyang (2010). "Forecasting with Mixed Frequencies". *Federal Reserve Bank of St. Louis Review*, November/December 2010, 92(6), pp. 521-36.

[4] Aruoba, S., Diebold, F. and C. Scotti (2009). "Real-Time Measurement of Business Conditions".*Journal of Business and Economic Statistics* 27, 4, pp. 417-27.

[5] Banbura M., Giannone, D. and L. Reichlin (2010) "Nowcasting". European Central Bank Working Paper Series, No 1275.

[6] Banbura M., Giannone, D., Modugno, M. and L. Reichlin (2013) "Now-casting and the real-time data flow", *Handbook of Economic Forecasting*, Elliot G. and A. Timmermann, eds., Chapter 4, Vol. 2, North-Holland.

[7] Bates, J.M. and C. W. J. Granger (1969), "The Combination of Forecasts", Operational Research Society, Vol. 20, No. 4., pp.451 - 468.

[8] Bell, V., Co, L., Stone, S. and G.Wallis (2014). "Nowcasting UK GDP Growth", Bank of England Quarterly Bulletin Q1.

[9] Boivin, J. and S. Ng (2006). "Are more data always better for factor analysis?". *Journal of Econometrics*, Vol. 132, issue 1, pp. 169-194.

[10] Caggiano, G., Kapetanios, G. and V. Labhard (2009). "Are more data always better for factor analysis? Results for the euro area, the six largest euro area countries and the UK". European Central Bank Working Paper Series, No 1051.

[11] Camacho, M., Perez-Quiros, G. and P. Poncela (2013). "Short-term forecasting for empirical economists. A survey of the recently proposed algorithms". Documentos de Trabajo N.º 1318, Banco de España.

[12] Clark, T. and K. West, (2007). "Approximately normal tests for equal predictive accuracy in nested models", *Journal of Econometrics*, Vol. 138, Issue 1, pp. 291-311.

[13] Clements, M. and D. Hendry (2006). "Forecasting with breaks" in Handbook of Economic Forecasting, Elliot, G., C. Granger and A. Timmermann, eds., Chapter 12, Vol. 1, North-Holland.

[14] Croushore, D. (2006). "Forecasting with real-time macroeconomic data", in *Handbook of Economic Forecasting*, Elliot, G., C. Granger and A. Timmermann, eds., Chapter 17, Vol. 1, North-Holland.

[15] D'Amato, L., Garegnani, L. and E. Blanco (2009). "Forecasting Inflation in Argentina: Individual Models or Forecast Pooling?", *Monetaria*, Centro de Estudios Monetarios Latinoamericanos, vol. 0(2), pp. 151-179, apr-jun.

[16] Diebold, F. and R.S. Mariano (1995). "Comparing Predictive Accuracy", *Journal of Business & Economic Statistics*, No.13, pp. 253-263.

[17] Drechsel, K. and L. Maurin (2008). "Flow of Conjuntural Information and Forecast of Euro Area Economic Activity", ECB WP No. 925, August.

[18] Evans, M. (2005). "Where are we now? Real-time Estimates of the Macro Economy", The International Journal of Central Banking, Vol. 1, 2, pp. 127-175.

[19] Giacomini, R. and B. Rossi, (2010). "Forecast Comparisons in Unstable Environments, with ", *Journal of Applied Econometrics* 25(4), April 2010, pp. 595-620.

[20] Giacomini, R. and B. Rossi, (2013). "Forecasting in macroeconomics", *Handbook of Research Methods and Applications in Empirical Macroeconomics*, Chapter 17, pp. 381-408, Edward Elgar Publishing.

[21] Giacomini, R. and H. White, (2004). "Tests of conditional predictive ability", *Econometrica*, Vol 74 No 6, pp. 1545-1578.

[22] Ghysels, E. (2013). "MATLAB Toolbox for Mixed Sampling Frequency Data Analysis Using MIDAS Regression Models." Available on MATLAB Central at http://www.mathworks.com/matlabcentral/fileexchange/45150-midas-regression.

[23] Ghysels E., Kvedaras V., Zemlys V. (2016). "Mixed Frequency Data Sampling Regression Models: The R Package midasr." R package version 0.6, URL https://CRAN.R-project.org/package=midasr.

[24] Foroni, C.and M. Marcellino (2013). "A Survey of Econometric Methods for Mixed-Frequency Data", EUI Working Paper ECO 2013/02.

[25] Giannone, D., L. Reichlin and D. Small (2005). "Nowcasting GDP and Inflation: The Real Time Informational Content of Macroeconomic Data Releases", *CEPR Discussion Papers* 5178, C.E.P.R. Discussion Papers.

[26] Giannone, D., L. Reichlin and D. Small (2008). "Nowcasting: The real-time informational content of macroeconomic data". *Journal of Monetary Economics* 55 (2008) 665– 676.

[27] Ghysels, E., Santa-Clara, P. and R. Valkanov (2004). "The MIDAS Touch: Mixed Data Sampling Regression Models", CIRANO Working Papers 2004s-20, CIRANO.

[28] Ghysels E., Kvedaras V., Zemlys V. (2016). "Mixed Frequency Data Sampling Regression Models: The R Package midasr." Journal of Statistical Software, Volume 72, Issue 4. https://www.jstatsoft.org/article/view/v072i04

[29] Granger, C. and R. Ramanathan (1984). "Improved methods of forecasting", *Journal of Forecasting*, Vol. 3, pp. 197-204.

[30] Hendry, D. and M. Clements, (2002). "Pooling of forecasts", *Econometrics Journal*, Vol. 5, pp. 1-26.

[31] Jansen W. J., X. Jin and J M.de Winter, (2016). "Forecasting and nowcasting real GDP: Comparing statistical models and subjective forecasts", International Journal of Forecasting 32, pp 411-436.

[32] Kitchen, J. and R. Monaco (2003). "Real-Time Forecasting in Practice", Business Economics, Department of the US Treasury, October.

[33] Marcellino, M. (2002). "Forecasting pooling for short time series of macroeconomic variables", *Oxford Bulletin of Economics and Statistics* No. 66, pp. 91-112.

[34] Pincheira, P. (2006). "Conditional evaluation of exchange rate predictive ability in long run regressions", Central Bank of Chile, Working Paper No. 378.

[35] Rünstler, G. and F. Sèdillot (2003). "Short-term estimates of Euro Area real GDP by means of monthly data", European Central Bank Working Paper No. 276.

[36] Stock, J. and M. Watson (2002a), "Macroeconomic Forecasting using diffusion indexes", *Journal of Business and Economic Statistics*, Vol. 20, pp. 147-162.

[37] Stock, J. and M. Watson (2006). "Forecasting with many predictors", in *Handbook of Economic Forecasting*, Elliot, G., C. Granger and A. Timmermann, eds., Chapter 10, Vol. 1, North-Holland.

[38] Timmermann A. (2006). "Forecast Combination", in *Handbook of Economic Forecasting*, Elliot, G., C. Granger and A. Timmermann, eds., Chapter 4, Vol. 1, North-Holland.

[39] Watson, M. (2001). "Macroeconomic Forecasting Using Many Predictors", in *Advances in Economics and Econometrics: Theory and Applications*, Eight World Congress, Vol. III, Chapter 3, Econometric Society.

[40] Stock, J. and Watson, M. (2016). "Dynamic Factor Models, Factor-Augmented Vector Autoregressions, and Structural Vector Autoregressions in Macroeconomics", in *Handbook of Macroeconomics*, John B. Taylor and Harald Uhlig, eds., Chapter 8, Vol. 2, pp. 415-525.

[41] West, K. (2006). "Forecast Evaluation", in *Handbook of Economic Forecasting*, Elliot, G., C. Granger and A. Timmermann, eds., Chapter 3, Vol. 1, North-Holland.

Annex 1: Complete Data Set

No	Source	Series	Group	
serie1	ADEFA	Automobile national production - units	1	hard
serie2	ADEFA	Automobile exports - units	1	hard
serie3	ADEFA	Automobile sales - units	1	hard
serie4	ADEFA	Automobile national sales - units	1	hard
serie5	AFCP	Portland cement production	1	hard
serie6	MECON	Ganancias (Total)	1	hard
serie7	MECON	Ganancias DGI	1	hard
serie8	MECON	Ganacias DGA	1	hard
serie9	MECON	Total Income revenues	1	hard
serie10	MECON	Income revenues DGI	1	hard
serie11	MERVAL	Income revenues DGA (customs)	1	prices
serie12	MERVAL	Total VAT revenues	1	prices
serie13	BCRA	VAT revenues DGI	1	prices
serie14	BCRA	Interest rate on Time Deposits - Private Banks	1	prices
serie15	CCA	Used Car Sales	1	hard
serie16	UTDT	Consumer Confidence Index - General - BSAS city	1	soft
serie17	UTDT	Consumer Confidence Index - General	1	soft
serie18	UTDT	ICC-DI	1	soft
serie19	UTDT	ICC-SM	1	soft
serie20	UTDT	ICC-SP	1	soft
serie21	UTDT	ICC-Condiciones Presentes	1	soft
serie22	UTDT	ICC-Expectativas	1	soft
serie23	CIS	Hierro Primario	1	hard
serie24	CIS	Acero Crudo	1	hard
serie25	CIS	Lam. Frío	1	hard
serie26	CIS	Lam. En caliente Total No Planos	1	hard
serie27	CIS	Lam. En caliente Planos	1	hard
serie28	FIEL	Industrial production index (IPI) - general level	2	hard
serie29	FIEL	IPI - nondurable consumer goods	2	hard
serie30	FIEL	IPI - durable consumer goods	2	hard

No	Source	Series	Group	
serie31	FIEL	IPI - intermediate goods	2	hard
serie32	FIEL	IPI - capital goods	2	hard
serie33	FIEL	IPI - food and beverages	2	hard
serie34	FIEL	IPI - cigarettes	2	hard
serie35	FIEL	IPI - textiles input	2	hard
serie36	FIEL	IPI - pulp and paper	2	hard
serie37	FIEL	IPI - fuels	2	hard
serie38	FIEL	IPI - chemicals and plastic	2	hard
serie39	FIEL	IPI - nonmetallic minerals	2	hard
serie40	FIEL	IPI - steel	2	hard
serie41	FIEL	IPI - metalworking	2	hard
serie42	FIEL	IPI - automobiles	2	hard
serie43	Gov. BSAS city - CABA	Gross Revenue Tax Collection - City of Buenos Aires	2	hard
serie44	Gov. BSAS Prov. (State)	Gross Revenue Tax Collection - Buenos Aires province	2	hard
serie46	CAME	Sales - General Level	1	hard
serie47	CAME	Sales - FOOD AND DRINKS	1	hard
serie48	CAME	Sales - BAZAAR AND GIFTS	1	hard
serie49	CAME	Sales - Bijouterie	1	hard
serie50	CAME	Sales - Shoes	1	hard
serie51	CAME	Sales - sports	1	hard
serie52	CAME	Sales - Home appliances	1	hard
serie53	CAME	Sales - Pharmacies	1	hard
serie54	CAME	Sales - Hardware store	1	hard
serie55	CAME	Sales - Candy and Soft Drinks	1	hard
serie56	CAME	Sales - Toy stores	1	hard
serie57	CAME	Sales - Leather Goods	1	hard
serie58	CAME	Sales - Electrical Supplies	1	hard
serie59	CAME	Sales - Construction materials	1	hard
serie60	CAME	Sales - Home furniture	1	hard

No	Source	Series	Group	
serie61	CAME	Sales - Office furniture	1	hard
serie62	CAME	Sales - Perfumery	1	hard
serie63	CAME	Sales - Textile - Clothing	1	hard
serie64	CAME	Sales - Textile - White	1	hard
serie65	CONSTRUYA	Construction Companies Activity Index	1	#N/A
serie66	CONSTRUYA	Construction Companies Activity Index SA	1	hard
serie67	INDEC	Exports - General Level	2	hard
serie68	INDEC	Exports - Q Primary Products	2	hard
serie69	INDEC	Exports - Q manufactures of agricultural origin	2	hard
serie70	INDEC	Exports - Q manufactures of industrial origin	2	hard
serie71	INDEC	Exports - Q Fuels and energy	2	hard
serie72	INDEC	Exports - P General level	2	hard
serie73	INDEC	Exports - P Primary Products	2	prices
serie74	INDEC	Exports - P manufactures of agricultural origin	2	prices
serie75	INDEC	Exports - P manufactures of industrial origin	2	prices
serie76	INDEC	Exports - P Fuels and energy	2	prices
serie77	INDEC	Imports - Q General level	2	hard
serie78	INDEC	Imports - Q capital goods	2	hard
serie79	INDEC	Imports - Q intermediate goods	2	hard
serie80	INDEC	Imports - Q Fuels and energy	2	hard
serie81	INDEC	Imports - Q Parts and Accessories	2	hard
serie82	INDEC	Imports - Q consumer goods	2	hard
serie83	INDEC	Imports - vehicles	2	hard
serie84	INDEC	Imports - P General level	2	prices
serie85	INDEC	Imports - P capital goods	2	prices
serie86	INDEC	Imports - P intermediate goods	2	prices
serie87	INDEC	Imports - P Fuel and energy	2	prices
serie88	INDEC	Imports - P Parts and Accessories	2	prices
serie89	INDEC	Imports - P consumer goods	2	prices
serie90	INDEC	Imports - P vehicles	2	prices

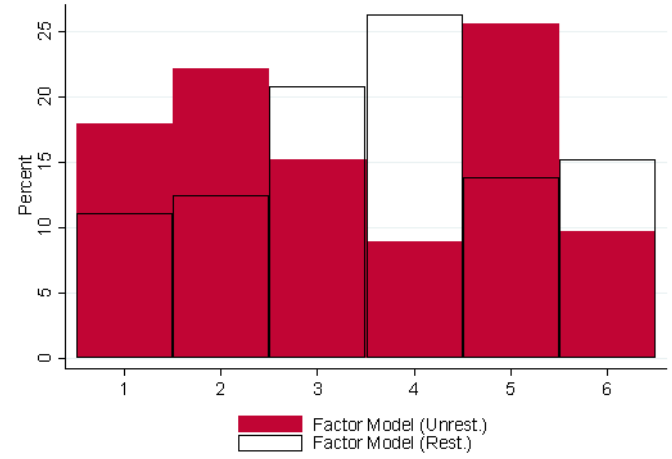
No	Source	Series	Group	
serie91	Ministerio de Agroindustria	Soybean milling	2	hard
serie92	Secretaría de Hacienda	Direct real investment + capital transfers to provinces	2	hard
serie93	Secretaría de Hacienda	Direct real investment	2	hard
serie94	Secretaría de Hacienda	Capital transfers to provinces	2	prices
serie95	Tendencias	Dismissals (1986 = 100)	1	soft
serie96	Tendencias	Suspensions (1986 = 100)	1	soft
serie97	EIL - Ministerio de Trabajo de la Nación	Net employment expectancy	2	soft
serie98	EIL - Ministerio de Trabajo de la Nación	Companies that searched for personnel	2	soft
serie99	BCRA	Multilateral nominal exchange rate index (Dec-15=100)	1	prices
serie100	BCRA	Personal Credits	1	prices
serie101	BCRA	Credit Cards	1	prices
serie102	BCRA	Personal + Cards	1	prices
serie103	GCBA	Vehicule Registrations BSAS city	2	hard
serie104	GCBA	Vehicule Registrations Argentina	2	hard
serie105	GCBA	Tolls (collection)	2	hard
serie106	GCBA	Tolls (vehicle ciculation)	2	hard
serie107	GCBA	Tolls (average vehicles)	2	hard
serie108	GCBA	Stamp duty-BSAS city	2	hard
serie109	GCBA	Passengers transported by rail (in thousands)	2	hard
serie110	Banco Central de BRASIL	Brazil Industrial production s.a.	2	hard
serie111	Banco Central de BRASIL	Brazil Industrial production	2	hard
serie112	Banco Central de BRASIL	Brazil Activity indicator s.o.	2	hard
serie113	Banco Central de BRASIL	Brazil Activity indicator s.a.	2	hard
serie114	Secretaria de energía	Asphalt (in tonnes)	2	hard
serie115	Colegio de escribanos Buenos Aires	BSAS city Scriptures	2	hard

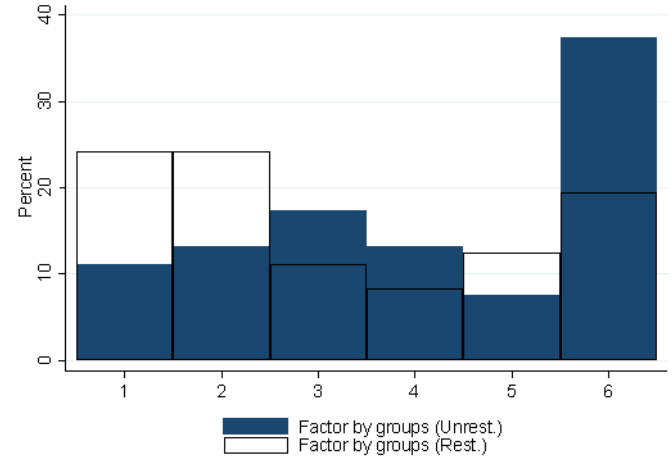
Selected series

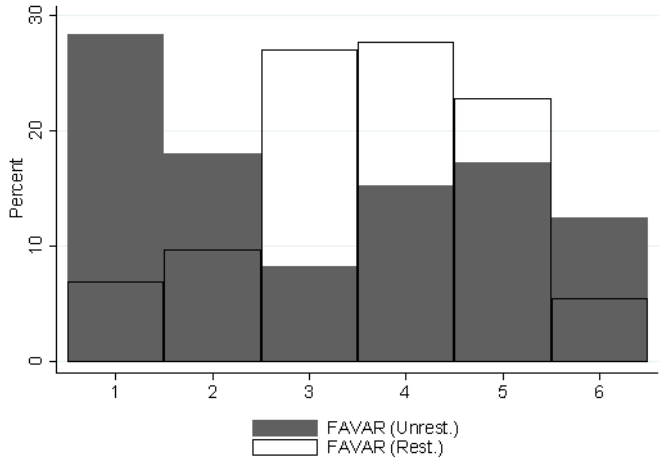
SeriesID	Frequency	Transformation	Units	Category
serie1	m	pch	units	Manufacturing
serie2	m	pch	units	Manufacturing
serie5	m	pch	units	Manufacturing
serie6	m	pch	$, Millons	Tax
serie15	m	pch	units	Retail and consumption
serie17	m	chg	Index	Surveys
serie20	m	chg	Index	Surveys
serie21	m	chg	Index	Surveys
serie26	m	pch	units	Manufacturing
serie30	m	pch	Index	Manufacturing
serie32	m	pch	Index	Manufacturing
serie42	m	pch	Index	Manufacturing
serie51	m	pch	Index	Retail and consumption
serie52	m	pch	Index	Retail and consumption
serie54	m	pch	Index	Retail and consumption
serie55	m	pch	Index	Retail and consumption
serie57	m	pch	Index	Retail and consumption
serie58	m	pch	Index	Retail and consumption
serie59	m	pch	Index	Retail and consumption
serie61	m	pch	Index	Retail and consumption
serie63	m	pch	Index	Retail and consumption
serie70	m	pch	Index	Housing and construction
serie75	m	pch	Index	International trade
serie77	m	pch	Index	International trade
serie79	m	pch	Index	International trade
serie81	m	pch	Index	International trade
serie82	m	pch	Index	International trade
serie97	m	chg	Index	Surveys
serie110	m	pch	Index	National account
serie113	m	pch	Index	National account

Annex 2: Histogram of relative positions for each model (full Sample)

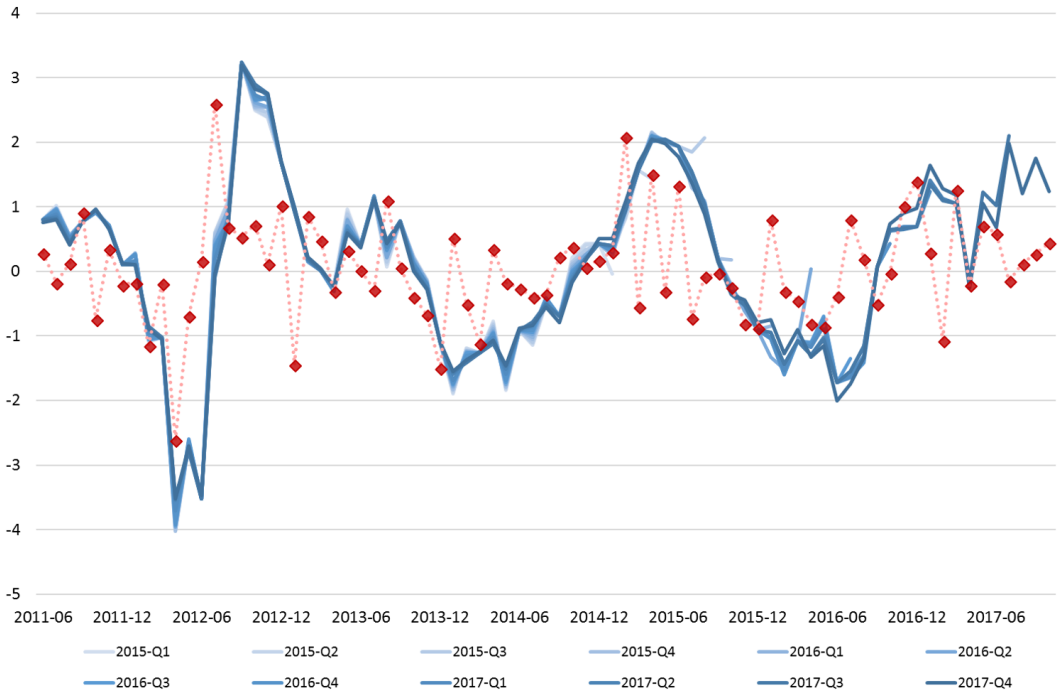
We compare the 6 models (Restricted and Unrestricted Factor Model, Factor by groups Model and FAVAR) and rank in decreasing order, i.e. 1 is the best model, 6 the worst.







Annex 3: Sequential Monthly GDP Estimations



A byproduct of the methodology is to obtain a monthly estimate of the GDP from the Kalman filter (i.e. a unobservable monthly component). The previous figure compares the sequential updates (blue/ish solid lines) with the latest monthly economic activity indicator or EMAE for its initials in Spanish (red dots). Overall, our estimations capture turning points reasonably. Since the official statistic office releases EMAE with a 2 month delay, there is also some informational gain from constructing a early monthly GDP estimation.