ASOCIACION ARGENTINA
DE ECONOMIA POLITICA

ANALES | ASOCIACION ARGENTINA DE ECONOMIA POLITICA

XLVI Reunión Anual

Noviembre de 2011

ISSN 1852-0022
ISBN 978-987-99570-9-7

NOWCASTING ECONOMIC ACTIVITY IN ARGENTINA WITH MANY PREDICTORS

D'Amato, Laura
Garegnani, Lorena
Blanco, Emilio

Nowcasting economic activity in Argentina with many predictors [*,†]

Laura D'Amato
BCRA, UBA and UNLP

Lorena Garegnani
BCRA and UNLP

Emilio Blanco
BCRA

August 2011

Abstract

We pool a large data set of business cycle indicators to produce *Nowcast* of contemporaneous GDP growth. We also conduct *Nowcast* using factors for a restricted subset of the indicators. Using an *AR(1)* benchmark to compare the forecasting performance of both *Nowcasts*, we conclude that only the *Nowcast* with pooling outperforms this univariate model. The Giacomini and White (2004) test is employed to evaluate the out of sample forecasting performance of the pooling compared to the *AR(1)*. In general, results indicate that a rich data set approach can provide valuable predictions about GDP behavior for the immediate future.

Keywords: Forecast pooling, Large dataset, Real time forecast, Factor Models
JEL classification: C22, C53, E17

[*] The opinions expressed in this work are those of the authors, and do not necessarily reflect the opinions of the Central Bank of Argentina or its authorities.

[†] We thank Hildegart Ahumada and Pablo Pincheira for valuable suggestions and comments.

1 Introduction

While real time assessment of the state of the economy as well as forecast of its future path are key for the conduct of monetary policy, the main source of information on economic activity are the national accounts, which are released on a quarterly basis.

Recent advances in the forecasting literature, focused on working in a rich-data environment could be very helpful to deal with this problem.[1] Here we will concentrate on two strategies developed to profit from the availability of a large number of business cycle indicators to improve forecasts: Factor models and forecast pooling (see Stock and Watson, 2006). Both have proved to deliver good results in terms of forecast accuracy.

Rich data sets can also be profited from at different frequencies during the quarter. In fact, a large data set of daily, weekly and monthly indicators are in general available to predict GDP within the quarter, what is known as *Nowcast* in literature. This approach is "real-time" because the estimate for current quarter GDP growth can be updated using the flow of conjunctural information as new data become available. Early forecast of GDP is particularly important, taking into account that official GDP figures for a specific quarter are released around 10 weeks after it ends.

Using a large set of daily, weekly and monthly business cycle indicators we construct a pooling and conduct a *Nowcast* exercise of GDP growth. In the case of data based contemporaneous GDP growth *Nowcast*, we assess the information content of these indicators in terms of the improvement they produce in forecast accuracy when sequentially added to the information set used to estimate current GDP growth. Individual estimations are combined using *R-squared* values of fitted models, which appears to be a plausible weighting method when prediction is based on estimated fitted values. We compare the performance of this combination against an *AR(1)* model used as a benchmark.

We additionally conduct a *Nowcast* exercise with factors using a subset of the selected business cycle indicators. We estimate factors through principal components techniques as introduced by Stock and Watson (2002a).

Using the Giacomini and White (2004) approach we also evaluate the out of sample predictive performance of the pooling to forecast GDP growth compared to the *AR(1)* as the natural benchmark. This test focuses on conditional predictive ability, comparing rival forecasting methods in terms of today´s accuracy to produce forecast for the near future.

The paper is organized as follows. Section 2 briefly describes the new developments in the forecasting literature related to working on a rich data environment. Our empirical approach and the results obtained from the *Nowcast* exercise are presented in section 3. We conduct an out of sample forecast exercise in section 4. Finally, section 5 concludes.

2 Methodology: Forecasting in a rich data environment

Causal econometric models often provide a satisfactory representation of the data-generating process (DGP) in terms of the behavior suggested by economic theory. These models do however tend to perform poorly when forecasting relevant time series, compared to simple devices.

[1] See in this respect Timmerman (2006).

One reason for this is that the latter tend to respond better to unanticipated changes in the data-generating process, given their intrinsically adaptive nature.

In recent years the forecasting literature has progressed in several directions in order to deal with these difficulties. Models employing a large number of predictors to forecast are now widely used. These models were developed in two avenues:

(i) *Forecast pooling*, which combines a considerable number of models using different weighting criteria.

(ii) *Factor models*, which make it possible to find summarized measures of the variability of a large number of relevant business cycle indicators.

In the first case, the path chosen aims to preserve the causal models and eventually to achieve better forecasts by expanding the group of predictors. In the second, a large set of business cycle indicators is considered, and by means of multivariate statistical techniques, a reduced number of factors -that explain a significant portion of their variability- underlying those series is extracted. Empirical evidence indicates that these variables add relevant information to explain the variability of the predicted variable.

2.1 Nowcast using pooling

Nowcast of a given economic indicator y_t implies conducting contemporaneous assessment of incoming information to produce continuous updates of forecast -as flows of conjuntural information become available. *Nowcast* uses a wide variety of x_t indicators and their bivariate relationships with y_t to predict it within the quarter.

It can be performed using bridge equations, linking monthly data released within the quarter with quarterly data. These individual-indicator forecasts are then aggregated using different weighting criteria to obtain an overall forecast of y_t. Individual autoregresive distributed lag models are estimated for each indicator and their fitted values are combined to produce a prediction of y_t for the current period. Although alternative procedures for combining individual bivariate forecasts are available, the use of weights based on in-sample relative explanatory power *(R^2)*, seems as a plausible way to produce *Nowcast*.

The *Nowcast* procedure using bridge equations and pooling works as follows: **(i)** selects the most recent data available by indicator, **(ii)** estimates the bivariate equation based on the last data available by indicator, **(iii)** produces a forecast by indicator and **(iv)** combines the individual forecasts according to their explanatory power. One of the benefits of this approach is that the regressions do not use forecasts of the independent variables.

As stressed in the literature, the combination of forecasts provides advantages at various levels:

(i) Forecast combinations provide diversification. Intuitively, when there is a quadratic loss function, even if one of the models outperforms another in terms of predictive power, by generating a lower loss, a linear combination could be preferable.[2]

[2] For a detailed view of the advantages of combining forecasts, see Hendry and Clements (2002), Marcellino (2002) and Timmermann (2006).

(ii) In the case of economies subject to structural changes, forecast combinations offer better predictions than individual models. In general, the speed at which models adapt to structural changes tends to differ. In such an instance, combination of models with different adaptability to changes could improve on individual models forecast.

(iii) Forecast combination could be seen as a way of making forecasts more robust in the face of specification bias and variable measurement errors in individual forecasts. For example, if two forecasts have different biases, in opposing directions, it is easy to see that a combination of both could generate an improvement in the forecast.

(iv) As stressed by Clements and Hendry (2006) forecast pooling can help dealing with structural breaks. In fact, they propose a battery of forecasting models that take into account break points in the mean and changes in deterministic trend.

The pooling or combination of forecasts implies combining two or more forecasts derived from models that use different predictors to produce a forecast. This technique was originally developed by Bates and Granger (1969), and the basic idea is as follows:[3]

Let $\{Y_{i,t+h}, i = 1, ..., n\}$ be a panel of n forecasts at time t for the $t + h$ period of a target variable Y. The combined forecast or forecasting pool will be given by the linear combination

$$Y_{t+h|t} = w_0 + \sum_{i=1}^{n} w_{it} Y_{i,t+h|t}$$

where w_{it} is the weight of the i^{th} forecast in period t.

Bates and Granger (1969) show that the weights that minimize the root mean squared forecast error (*RMSFE*) are given by the projection to the population of $Y_{t+h|t}$ in a constant and the individual forecasts. Frequently the constant is omitted, and by imposing $\sum_{i=1}^{n} w_{it} = 1$ it is determined that if each of the forecasts is unbiased, so is $Y_{t+h|t}$. As long as none of the forecasts is generated by the real model, the optimal combination of forecasts spreads the weight over a multiple combination of forecasts. The minimum *RMSFE* combining those forecasts will be variable over time if the variance and covariance matrixes for $\left(Y_{t+h|t}, \{Y_{i,t+h|t}\}\right)$ change over time.

In practice, optimal weightings are not viable because the variance and covariance matrixes are unknown. Granger and Ramanathan (1984) propose estimating weights using minimum least squares or restricted least squares, if $w_0 = 0$ and $\sum_{i=1}^{n} w_{it} = 1$ is imposed, although if n is large it is expected that estimates will perform poorly, simply because by estimating a large number of parameters, uncertainty is introduced into the sample. If n is proportionate to the size of the sample, the OLS estimator is not consistent, and the combinations that use it are not asymptotically optimum. For this reason, research into combination or pooling of forecasts has concentrated on imposing greater structure to the combination of forecasts. Among several weighting techniques we use the following two:

(i) *Weights based on in sample relative explanatory power (R^2)*: which combines forecast according to the strength of the estimated past relationship between each indicator and GDP growth.[4]

---

[3] A detailed description of forecast pooling techniques and the principal developments contained in this literature can be found in Stock and Watson (2006), and in even greater detail in Timmerman (2006).

[4] See Kitchen and Monaco (2003).

$$w_i = \frac{R_i^2}{\sum_{j=1}^{n} R_j^2}$$

where $j = 1, ...n$ are the monthly indicators considered to forecast GDP growth.

(ii) *Weights based on out of sample performance (RMSFE)*: In this case the combined forecast is constructed assigning weights which are inversely related to individual forecast *RMSFE*

$$w_{it} = \frac{m_{it}^{-1}}{\sum_{j=1}^{n} m_{jt}^{-1}}, \quad where\ m_{it} = \sqrt{\frac{\sum_{t=T+1}^{T+h} (\hat{y}_{i,t} - y_t)^2}{h}}$$

2.2 Nowcast using factors

Nowcast can also 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 underlying the use of Dynamic Factor Models is that 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 for n observables 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.

Given a vector of n stationary monthly indicators time series $x_t = (x_{1t}, ...x_{nt})\acute{}$, with $t = 1, ...T$, the vector for n observables in the cycle can be explained by the distributed lags of q common latent factors plus n idiosyncratic disturbances which could eventually be serially correlated or correlated among the $i's$

$$X_{it} = \lambda_i(L)\acute{} f_t + u_{it} \tag{1}$$

Where f_t is a vector $q\times 1$ of unobserved factors, λ is a $q\times 1$ vector lag polynomial of *dynamic factor loadings* and the u_{it} are the idiosyncratic disturbances that are assumed to be uncorrelated with the factors in all leads and lags, that is to say $E(f_t u_{it}) = 0 \ \forall \ i, s.$

The objective is therefore to estimate $E(y_t \mid X_t)$ modeling y_t according to

$$y_t = \beta(L)\acute{} f_t + \varepsilon_t \tag{2}$$

If the lag polynomials $\lambda_i\,(L)$ in (1) and $\beta\,(L)$ in(2) are of finite order p, 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 could use the bridge equation

$$\widehat{y_t}^Q = \beta(L)\acute{} f_t^Q$$

to produce *Nowcast* of GDP within the quarter.

3 The nowcasting exercise

3.1 Our empirical approach

The data comprises a broad set of 55 economic indicators ranging from financial indicators to tax collection data, business surveys, disaggregated data on industrial production, use of energy at the industry level and cars sales.[5] The sample used to estimate models is 1993:QI-2004:QI. We perform *Nowcast* and out of sample forecast for the period 2004:Q2-2007:QIV.

Series were seasonally adjusted when needed, de-trended or differenciated to make them stationary and finally log transformed.

In order to produce *Nowcast*, we estimate autoregressive distributed bivariate models with four lags of GDP for each of the corresponding business cycle indicators.

$$y_t = \alpha_0 + \sum_{i=1}^{4} \alpha_i y_{t-i} + \sum_{i=0}^{4} \beta_i x_{jt-i} + \epsilon_t$$

Where y is real GDP growth and x_j corresponds to the j^{th} indicator calculated at a quarterly rate as to make it homogeneous with output.

Following Drechsel and Maurin (2008) we estimate simple models regressing GDP growth on individual indicators. This procedure helps to reduce the problem of over-fitting and poor forecast performance.

Models were specified as to ensure white noise, homoskedastic and normally distributed residuals. Although very simple, models fit the data very well.[6] This is a promising property of models estimated for forecasting purposes, since it is highly probable that combining them would produce good out of sample forecast.

The *Nowcast* procedure is based on updating predictions according to incoming information. Given the diversity in the publication lags of the different indicators, the series are merged into three groups and converted to a quarterly basis to sequentially update the prediction of GDP for the current quarter.

The Argentine Institute of Statistics releases the official figures of quarterly GDP 10 weeks after the end of the respective quarter. *Figure 1* describes the sequential updating scheme of data releases of both, monthly indicators used to predict and quarterly GDP. As can be seen from the figure, six weeks after the beginning of the quarter a first *Nowcast* release for the quarterly GDP is available. This *Nowcast* is afterwards updated with the incoming information every fortnight.

[5] See Table A.1. in *Appendix I* for details.

[6] See Table A.2. in *Appendix I* for details on models.

Figure 1: Sequential updating scheme of data releases for monthly indicators and GDP

Date	15/02/2007	28/02/2007	15/03/2007	31/03/2007	15/04/2007	30/04/2007	15/05/2007	31/05/2007	15/06/2007
Available Data	Group 1 (jan obs) Group 2 (dec obs) Group 3 (dec obs)	Group 1 (jan obs) Group 2 (jan obs) Group 3 (dec obs)	Group 1 (feb obs) Group 2 (jan obs) Group 3 (jan obs)	Group 1 (feb obs) Group 2 (feb obs) Group 3 (jan obs)	Group 1 (mar obs) Group 2 (feb obs) Group 3 (feb obs)	Group 1 (mar obs) Group 2 (mar obs) Group 3 (feb obs)	Group 1 (apr obs) Group 2 (mar obs) Group 3 (mar obs)	Group 1 (apr obs) Group 2 (apr obs) Group 3 (mar obs)	Group 1 (may obs) Group 2 (apr obs) Group 3 (apr obs)
Nowcast	I 2007	I 2007	I 2007	I 2007	I 2007	I 2007	I 2007 II 2007	II 2007	II 2007
Official releases									**Fist Quarter official release**

3.2 Results

In *Table 1* we present the sequentially updated predicted values of GDP growth and their performance measured by the absolute value of the difference between actual and fitted values. It can be seen from the table that *Nowcast* performs exceptionally well for every quarter. It can also be noticed that it is not clear that performance improves with the addition of information. In fact, the prediction for the first month outperforms the prediction using the complete set of information for the current quarter in some quarters. The set of variables available at the end of the first month, which is used to produce this forecast includes monetary and financial indicators such as interest rates, stock prices, money aggregates, as well as tax revenues, automobile sales, steel and portland cement production, and energy demand, among others.

Table 1: Nowcast performance

Sequential Updating of current GDP growth predictions								
Date	Actual	15 days	1 month	45 days	2 months	75 days	3 months	105 days
2004-Q2	0.01092	0.01452	0.01458	0.01717	0.01711	0.01802	0.01819	0.01698
2004-Q3	0.02306	0.01092	0.00924	0.00938	0.01025	0.01044	0.01133	0.01150
2004-Q4	0.02668	0.02638	0.02649	0.02311	0.02314	0.02411	0.02411	0.02362
2005-Q1	0.02240	0.01457	0.01818	0.01600	0.01538	0.01683	0.01717	0.01871
2005-Q2	0.02375	0.01506	0.01650	0.01439	0.01418	0.01478	0.01468	0.01588
2005-Q3	0.01914	0.01960	0.02174	0.02051	0.01922	0.01781	0.01885	0.01893
2005-Q4	0.01906	0.01649	0.01931	0.01766	0.01752	0.01605	0.01577	0.01544
2006-Q1	0.02146	0.00991	0.01264	0.01435	0.01411	0.01782	0.01854	0.01759
2006-Q2	0.01952	0.01934	0.01934	0.01988	0.02005	0.02006	0.01985	0.01882
2006-Q3	0.02372	0.02048	0.01794	0.01706	0.01757	0.01664	0.01684	0.01661
2006-Q4	0.01741	0.02385	0.02036	0.01934	0.02005	0.01869	0.01870	0.01870
2007-Q1	0.01716	0.01770	0.01685	0.01287	0.01396	0.01341	0.01778	0.01778
2007-Q2	0.02322	0.01807	0.02080	0.01660	0.01643	0.01782	0.01779	0.01714
2007-Q3	0.01927	0.02222	0.01897	0.01632	0.01756	0.02112	0.02101	0.01929
2007-Q4	0.02379	0.02301	0.02296	0.02369	0.02387	0.02127	0.02130	0.02023
Sequential Updating: Evolution of predictive preformance								
Date	Actual	15 days	1 month	45 days	2 months	75 days	3 months	105 days
2004-Q2	0.01092	0.00361	0.00366	0.00625	0.00619	0.00711	0.00728	0.00607
2004-Q3	0.02306	0.01214	0.01381	0.01367	0.01281	0.01262	0.01173	0.01156
2004-Q4	0.02668	0.00031	0.00019	0.00358	0.00355	0.00258	0.00257	0.00307
2005-Q1	0.02240	0.00783	0.00422	0.00640	0.00702	0.00558	0.00524	0.00370
2005-Q2	0.02375	0.00869	0.00726	0.00937	0.00958	0.00897	0.00908	0.00788
2005-Q3	0.01914	0.00046	0.00260	0.00137	0.00008	0.00133	0.00029	0.00021
2005-Q4	0.01906	0.00256	0.00025	0.00140	0.00153	0.00301	0.00329	0.00362
2006-Q1	0.02146	0.01154	0.00881	0.00711	0.00735	0.00363	0.00292	0.00386
2006-Q2	0.01952	0.00018	0.00018	0.00037	0.00053	0.00054	0.00034	0.00069
2006-Q3	0.02372	0.00323	0.00577	0.00665	0.00615	0.00708	0.00688	0.00710
2006-Q4	0.01741	0.00644	0.00295	0.00193	0.00264	0.00128	0.00129	0.00129
2007-Q1	0.01716	0.00054	0.00031	0.00429	0.00320	0.00375	0.00062	0.00062
2007-Q2	0.02322	0.00515	0.00242	0.00661	0.00679	0.00539	0.00542	0.00607
2007-Q3	0.01927	0.00294	0.00030	0.00295	0.00171	0.00185	0.00173	0.00002
2007-Q4	0.02379	0.00079	0.00084	0.00011	0.00008	0.00253	0.00250	0.00357

In order to evaluate the predictive performance of *Nowcast* relative to a benchmark we compare the three months *Nowcast* estimation (including the complete set of information) with one quarter ahead forecast of an *AR(1)* model of GDP growth for the same quarter.

Table 2: Nowcast relative to benchmark

Date	Actual	*Forecast*		*Relative Forecasting Performance*	
		AR (1)	Nowcast	Actual-AR	Actual-Now
2004-Q2	0.01092	0.01899	0.01698	-0.00807	**-0.00607**
2004-Q3	0.02306	0.00834	0.01150	0.01472	**0.01156**
2004-Q4	0.02668	0.01698	0.02362	0.00971	**0.00307**
2005-Q1	0.02240	0.01984	0.01871	**0.00256**	0.00370
2005-Q2	0.02375	0.01699	0.01588	**0.00676**	0.00788
2005-Q3	0.01914	0.01816	0.01893	0.00098	**0.00021**
2005-Q4	0.01906	0.01497	0.01544	0.00409	**0.00362**
2006-Q1	0.02146	0.01502	0.01759	0.00643	**0.00386**
2006-Q2	0.01952	0.01689	0.01882	0.00263	**0.00069**
2006-Q3	0.02372	0.01559	0.01661	0.00812	**0.00710**
2006-Q4	0.01741	0.01879	0.01870	-0.00138	**-0.00129**
2007-Q1	0.01716	0.01427	0.01778	0.00289	**-0.00062**
2007-Q2	0.02322	0.01356	0.01714	0.00966	**0.00607**
2007-Q3	0.01927	0.01894	0.01929	0.00033	**-0.00002**
2007-Q4	0.02379	0.01786	0.02023	0.00593	**0.00357**

The results, shown in *Table 2*, indicate that the nowcast outperforms the *AR(1)* for 13 of the 15 quarters used to compare the predictive accuracy of the *Nowcast*. This result is strong since the *AR(1)* is the model usually considered as a benchmark for short term forecast, given its outperforming accuracy.

3.3 Testing for equal predictive ability

As usual in forecast evaluation,[7] we compare the predictive accuracy of the two previous nested models: The *AR(1)*, which is the parsimonious null model and a larger model, in our case the *Nowcast,* that nests the first one. Under the null the autoregresive model is the data generating process (*DGP*). Thus, the larger model is supposed to introduce noise because its forecast requires the estimation of population parameters that are supposed to be zero under the null. Given that, the *MSFE* of the *AR(1)* model is expected to be smaller than that of the *Nowcast*, Clark and West (2007) construct a test of equal predictive accuracy in nested models that takes into account this noise, adjusting the *MSFE* of the larger model.

The test proposed by Clark and West evaluates the predictive accuracy of the parsimonious model (in our case the *AR(1)*) with that of the larger model that nests it (the *Nowcast*).

The null hypothesis is that model 1 (*AR(1)*) has the same predictive accuracy than model 2 (the *Nowcast*). The alternative hypothesis is that model 2 has better predictive accuracy than model 1 (one sided test).

The test compares the *MSFE* of model 1 with the *MSFE* of model 2 adjusted for the squared difference of the forecasts of the two models.

$$\widehat{f}_{t+1} = (y_{t+1} - \widehat{y}_{1t,t+1})^2 - \left[(y_{t+1} - \widehat{y}_{2t,t+1})^2 - (\widehat{y}_{1t,t+1} - \widehat{y}_{2t,t+1})^2\right]$$

In practice, the test consists on regressing $\widehat{f}_{t+1}$ on a constant and evaluating its significance using the t statistic for the null of a 0 coefficient. Rejecting the null implies that the *MSFE* of the parsimonious model is significantly higher than that of the larger model.

Considering the sample period $2004-2007$ to compare the predictive accuracy of the two models we can reject the null of equal predictive accuracy, indicating that the *MSFE* of the *Nowcast*

[7] See West (2006).

is significantly lower than that of the *AR(1)* model. The value of the t statistic is 5.985 with a p value of 0.000, rejecting the null at the 1% level, what indicates that the *Nowcast* outperforms the benchmark.

3.4 Comparing forecasting performance of Nowcast using pooling with Nowcast using factors

As an additional way of evaluating the performance of the *Nowcast* using pooling, we conduct a simple *Nowcast* exercise using factors. We followed the methodology proposed by Stock and Watson (2002), as described in section 2.2, and estimated factors using principal components analysis. To solve the problem of dealing with an unbalanced panel of monthly indicators, we forecasted missing values using univariate autoregressive models. Factors were obtained from a selected subset of 17 series included in the group of business cycle indicators used to produce the *Nowcast* with the pooling of bridge equations. The reason to restrict the number of time series was that factors estimated with the complete set of business cycle indicators explained a small portion of the variance of the multivariate set of business cycle indicators. As shown by Boivin and Ng (2006) this outcome could be due to the fact that idiosyncratic errors in equation (1) are cross-correlated or because a subset of the time series considered are "noisy". In this case Boivin and Ng show that restricting the data set to estimate factors could be better. The 17 series used to construct the factors where selected according to a simple R^2 criteria. Out of the 55 series available, we kept those that presented a R^2 superior to 0.75 in an OLS regression for the period 1993Q1-2004Q1.[8]

Table 3 shows the performance of the Nowcast using factors.

[8] See Appendix I for a description of the time series included in the subset.

Table 3: Nowcast with factors performance

Sequential Updating of current GDP growth predictions with factors							
Date	Actual	15 days	1 month	45 days	2 months	75 days	3 months
2004-Q2	0.01092	0.01318	0.01119	0.03634	0.03851	0.03635	0.03639
2004-Q3	0.02306	0.01338	0.01250	0.02329	0.02363	0.02337	0.02364
2004-Q4	0.02668	0.02381	0.02364	0.02501	0.02534	0.02270	0.02263
2005-Q1	0.02240	0.00960	0.01041	0.00434	0.00418	0.01511	0.01486
2005-Q2	0.02375	0.01147	0.01099	-0.00779	-0.00796	-0.00034	-0.00033
2005-Q3	0.01914	0.00795	0.00809	0.03054	0.03052	0.02338	0.02336
2005-Q4	0.01906	0.02378	0.02384	0.03657	0.03642	0.02184	0.02192
2006-Q1	0.02146	0.00041	0.00037	-0.05577	-0.05639	0.03403	0.03378
2006-Q2	0.01952	0.01337	0.01348	0.03561	0.03566	0.03398	0.03377
2006-Q3	0.02372	0.02975	0.03035	0.02912	0.02910	0.02254	0.02242
2006-Q4	0.01741	0.02841	0.02854	0.03619	0.03629	0.03239	0.03243
2007-Q1	0.01716	0.02784	0.00816	-0.01819	-0.01742	-0.01992	0.02929
2007-Q2	0.02322	0.01209	0.01254	0.00751	0.00751	0.02347	0.02337
2007-Q3	0.01927	0.01452	0.01446	-0.01234	-0.01308	0.03381	0.03360
2007-Q4	0.02379	0.01304	0.02089	0.07484	0.07447	0.03758	0.03768
Sequential Updating: Evolution of predictive preformance							
Date	Actual	15 days	1 month	45 days	2 months	75 days	3 months
2004-Q2	0.01092	0.00226	0.00028	0.02543	0.02759	0.02544	0.02548
2004-Q3	0.02306	0.00968	0.01056	0.00023	0.00058	0.00031	0.00058
2004-Q4	0.02668	0.00288	0.00304	0.00168	0.00135	0.00398	0.00405
2005-Q1	0.02240	0.01281	0.01200	0.01806	0.01822	0.00729	0.00754
2005-Q2	0.02375	0.01229	0.01276	0.03154	0.03171	0.02409	0.02408
2005-Q3	0.01914	0.01119	0.01105	0.01140	0.01138	0.00424	0.00422
2005-Q4	0.01906	0.00472	0.00478	0.01752	0.01736	0.00278	0.00286
2006-Q1	0.02146	0.02104	0.02109	0.07722	0.07784	0.01257	0.01233
2006-Q2	0.01952	0.00615	0.00604	0.01610	0.01614	0.01446	0.01425
2006-Q3	0.02372	0.00604	0.00663	0.00541	0.00538	0.00117	0.00130
2006-Q4	0.01741	0.01100	0.01113	0.01878	0.01888	0.01498	0.01502
2007-Q1	0.01716	0.01068	0.00900	0.03535	0.03458	0.03708	0.01213
2007-Q2	0.02322	0.01113	0.01068	0.01570	0.01570	0.00025	0.00016
2007-Q3	0.01927	0.00475	0.00481	0.03161	0.03235	0.01454	0.01433
2007-Q4	0.02379	0.01075	0.00291	0.05105	0.05068	0.01378	0.01388

It can be noticed from that the *Nowcast* with factors performs quite well and again, it is not clear that adding information improves predictive accuracy.

We additionally perform a *Nowcast* based on a pooling of the 17 selected series used to construct the factors.

A comparison of the predictive performance of the *Nowcast* pooling with the complete and restricted sets and Nowcast with factors with a forecast using an *AR(1)* are presented in *Table 4*.

Table 4: Nowcast with pooling and factors relative to benchmark

Date	Actual	*Forecast*				*Relative Forecasting Performance*			
		AR (1)	Nowcast with Pooling	Nowcast with Factors	Nowcast with Pooling 17 series	Actual-AR	Actual-Now Pooling	Actual-Now Factors	Actual-Nowcast with Pooling 17 series
2004-Q2	0.01092	0.01899	0.01698	0.03639	0.02057	-0.00807	**-0.00607**	-0.02548	-0.00965
2004-Q3	0.02306	0.00834	0.01150	0.02364	0.01219	0.01472	0.01156	**-0.00058**	0.01087
2004-Q4	0.02668	0.01698	0.02362	0.02263	0.02670	0.00971	0.00307	0.00405	**-0.00002**
2005-Q1	0.02240	0.01984	0.01871	0.01486	0.01883	**0.00256**	0.00370	0.00754	0.00357
2005-Q2	0.02375	0.01699	0.01588	-0.00033	0.01273	**0.00676**	0.00788	0.02408	0.01102
2005-Q3	0.01914	0.01816	0.01893	0.02336	0.01895	0.00098	0.00021	-0.00422	**0.00019**
2005-Q4	0.01906	0.01497	0.01544	0.02192	0.01570	0.00409	0.00362	**-0.00286**	0.00335
2006-Q1	0.02146	0.01502	0.01759	0.03378	0.02344	0.00643	0.00386	-0.01233	**-0.00198**
2006-Q2	0.01952	0.01689	0.01882	0.03377	0.02161	0.00263	**0.00069**	-0.01425	-0.00210
2006-Q3	0.02372	0.01559	0.01661	0.02242	0.01672	0.00812	0.00710	**0.00130**	0.00699
2006-Q4	0.01741	0.01879	0.01870	0.03243	0.02066	-0.00138	**-0.00129**	-0.01502	-0.00325
2007-Q1	0.01716	0.01427	0.01778	0.02929	0.01811	0.00289	**-0.00062**	-0.01213	-0.00095
2007-Q2	0.02322	0.01356	0.01714	0.02337	0.01629	0.00966	0.00607	**-0.00016**	0.00692
2007-Q3	0.01927	0.01894	0.01929	0.03360	0.02317	0.00033	**-0.00002**	-0.01433	-0.00390
2007-Q4	0.02379	0.01786	0.02023	0.03768	0.02660	0.00593	0.00357	-0.01388	**-0.00281**

From *Table 4* it seems that all *Nowcast* models outperform the *AR(1)*. From the results of subsection 3.3. we know that the *Nowcast* with pooling (considering all series) outperforms the *AR(1)*.

Thus, in order to test for equal predictive accuracy of the *Nowcast* with factors relative to the *AR (1)* benchmark model, we also perform the Clark and West. Results show that the null hypothesis of equal predictive accuracy cannot be rejected at the 1% level. The value of the t statistic is 1.033 with a p value of 0.3042, giving evidence that the *Nowcast* with factors does not outperform the benchmark, contrary to what we find for the *Nowcast* with pooling.

When we conduct this test for the *Nowcast* with pooling (17 series) the value of the t statistic is 4.812, with a p value of 0.0000 indicating that this model outperforms the *AR(1)*.

These results suggest that the use of high frequency within the quarter information by means of a pooling is valuable to generate predictions of GDP growth.

4 Out of sample forecast of GDP growth

We also evaluate the out of sample forecasting performance of the pooling relative to that of the autoregressive GDP growth model used as benchmark. For this we conduct an out of sample forecast for 1 to 3 quarters horizon for the period 1999Q1-2009Q2.

In this case we compare the predictive accuracy of the two previous forecasting methods: The *AR(1)* and the *pooling,* using the methodology developed by Giacomini and White (2004).[9] The Giacomini and White approach differs from that followed by previous accuracy evaluation methodologies, as those proposed by Dieblod and Mariano (1995) and West (2003) in what it is based on conditional rather than unconditional expectations, as it is the case of the tests proposed by the latter. 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 that testing the validity of a theoretical model.

The test has many advantages: **(i)** it captures the effect of estimation uncertainty on relative forecast performance, **(ii)** is useful for forecasts based on both nested and nonnested models, **(iii)** allows the forecasts to be produced by general estimation methods, and **(iv)** 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 out of sample 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 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:

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

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

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 heteroksedasticity and autocorrelation.

The results of applying the Giacomini and White procedure to evaluate the forecasting performance of the pooling relative to the *AR(1)* are shown in *Table 5*. They show that the pooling of forecasts using the business cycle indicators outperforms the *AR(1)* at the very short horizons of one and two quarters. Thus, it seems that rich data set can provide valuable predictions about GDP behavior in the immediate future.

Table 5: Results of the Giacomini and White test

Sample: 1999Q1 - 2009Q2	
1 quarter ahead	-0.000189
	*(-3.41245)**
2 quarter ahead	-0.00033
	*(-1.959476)***
3 quarter ahead	-0.000391
	(-1.671884)

* Significant at 1%
** Significant at 6%
(.) Student t-test

5 Conclusions

While real time assessment of economic activity is crucial to evaluate the presence of inflationary pressures for the purpose of monetary policy decision making, GDP figures are produced in a quarterly basis and released with a 10 week lag.

Recent advances in the forecasting literature, focused on working in a rich-data environment have developed strategies to profit from the availability of a large number of business cycle indicators to improve forecast through the use of factor models and forecast pooling. Both have proved to deliver good results in terms of forecast accuracy.

In terms of producing real time forecast, pooling has the advantage of being flexible to sequentially update forecast at the time new information is released.

Using a large set of daily and monthly business cycle indicators we construct a pooling and conduct *Nowcast* with pooling of Argentina's GDP growth. We also develop a *Nowcast* exercise using factors from a restricted subset of business cycle indicators.

In the case of data based *Nowcast* with pooling of contemporaneous GDP growth we assess the information content of these indicators in terms of the improvement they produce in predictive accuracy when they are sequentially added to the information set used to estimate current GDP growth. In this case individual estimations are combined using *Rsquared* values of fitted models, what appears to be a plausible weighting scheme when prediction is based on estimated fitted values.

The results show that the *Nowcast* with pooling performs well, although it is not clear that the performance improves with the addition of information.

In order to evaluate the predictive performance of *Nowcast* with pooling relative to a benchmark we compare the three months *Nowcast* estimation with a forecast one quarter ahead of an *AR(1)* model for the GDP growth. The results indicate that the *Nowcast* with pooling outperforms the *AR(1)* model for 13 of 15 quarters considered. When we test the differences in predictive accuracy between both models we confirm that the *Nowcast* with pooling has better accuracy. We also evaluate the predictive performance of a factor model and a pooling using a subset of the complete set of business cycle indicators. In this case the predictive accuracy of the *Nowcast* with factors is not statistically different from that of the *AR(1)* while the *Nowcast* pooling with the restricted subset does outperform the *AR(1)*. Thus, *Nowcast* with pooling using both the complete dataset and a subset of incoming information within the quarter improve forecast accuracy of an *AR(1)*.

Additionally, we conduct an out of sample forecast for different horizons using the pooling of forecast and find that it outperforms the *AR(1)* for the very short run horizons.

These results for nowcasting and forecasting using a large data set are promising, since the *AR(1)* is the usual model chosen in the literature for short term forecasting.

The *Nowcast* methodology has a potentially broad application to any macro or goal variable of interest. It also represents a potentially valuable approach to provide timely information for policy decision making.

References

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

[2] Bates, J. and C. Granger (1969). "The combination of forecasts", *Operations Research Quarterly,* No. 20, pp. 451-468.

[3] 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.

[4] 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.

[5] 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.

[6] D'Amato, L., L. Garegnani and E. Blanco (2008). "Forecasting Inflation in Argentina: Individual Models or Forecast Pooling?", BCRA working Paper No. 35.

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

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

[9] Giacomini, R. and H. White, (2004). "Tests of conditional predictive ability", *Econometrica*, Vol 74 N^o 6, 1545-1578.

[10] Giannone, D., Reichlin, L. 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.

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

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

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

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

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

[16] 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.

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

[18] 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.

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

[20] 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.

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

Appendix I

The series used were seasonally adjusted (when necessary) using the X-12 ARIMA program, and were subsequently standardized either by differentiating them (*diff*) or by subtracting a linear deterministic trend (*trend*).[10] Table A1 presents the complete series.

Table A.1.: Data Set

Series	Release	Stationary
Group 1: 15 days delay		
Autobile national production - units	monthly	dif
Autobile exports - units	monthly	dif
Autobile sales - units	monthly	dif
Autobile national sales - units	monthly	dif
Portland cement production - thousands of tons	monthly	dif
Income Revenues	monthly	tend
Income Revenues - DGI	monthly	tend
Income Revenues - DGA	monthly	dif
VAT revenue	monthly	tend
VAT revenue -DGI	monthly	dif
MERVAL stock market index - monthly average*	daily	dif
MERVAL stock market index - at month-end	monthly	dif
Group 2: 1 month delay		
Steel rods for concrete production - tons	monthly	dif
Raw steel production - thousands of tons	monthly	dif
Cold rolled steel production - thousands of tons	monthly	dif
Hot rolled non-flat steel - thousands of tons	monthly	dif
Flat hot rolled steel - thousands of tons	monthly	dif
Energy demand sales - GWh	monthly	dif
Private M2*	daily	tend
Nominal interest rate - 30-59 days - private banks*	daily	dif
Group 3: 2 months delay		
Industrial Survey - industry stock levels manufacturing	monthly	dif
Industrial Survey - non-durable cons. goods stock levels	monthly	dif
Industrial Survey - consumer durables stock levels	monthly	dif
Industrial Survey - capital gods stock levels	monthly	dif
Industrial Survey - intermediate goods stock levels	monthly	dif
Industrial Survey -outlook manufacturing industry	monthly	dif
Industrial Survey -outlook non-durable cons. Goods	monthly	dif
Industrial Survey -outlook consumer durables	monthly	dif

[10] Standard ADF unit root tests were conducted.

Group 3: 2 months delay (cont.)		
Industrial Survey -outlook capital goods	monthly	dif
Industrial Survey -outlook intermediate goods	monthly	dif
Industrial Survey - general situation manufacturing industry	monthly	dif
Industrial Survey - general situation non-durable consumer goods	monthly	dif
Industrial Survey - general situation consumer durables	monthly	dif
Industrial Survey - general situation capital goods	monthly	dif
Industrial Survey - general situation intermediate goods	monthly	dif
Industrial Survey - manufacturing industry demand trend	monthly	dif
Industrial Survey - non-durable cons. goods demand trend	monthly	dif
Industrial Survey - consumer durables demand trend	monthly	dif
Industrial Survey - capital gods demand trend	monthly	dif
Industrial Survey - intermediate goods demand trend	monthly	dif
Industrial production index (IPI) - general level	monthly	dif
IPI - non-durable consumer goods	monthly	dif
IPI - durable consumer goods	monthly	dif
IPI - intermediate goods	monthly	dif
IPI - capital goods	monthly	dif
IPI - food and beverages	monthly	dif
IPI - cigarettes	monthly	dif
IPI - textiles input	monthly	dif
IPI - pulp and paper	monthly	dif
IPI - fuels	monthly	dif
IPI - chemicals and plastics	monthly	dif
IPI - non-metallic minerals	monthly	dif
IPI - steel	monthly	dif
IPI - metalworking	monthly	dif
IPI - automobiles	monthly	dif

*quarterly figures are obtained from averaging daily data

Table A.2.: Summary of models

Model	Dummies included	R2
variable 1	D1995M1 - D1995M2 - D2000M1 - D2001M3 - D2001M4	0.8185
variable 2		0.6493
variable 3	D2001M3 - D2001M4 - D2002M1	0.7406
variable 4		0.6042
variable 5	D1999M2	0.7487
variable 6		0.6328
variable 7	D1995M2 - D2000M1	0.7507
variable 8		0.6086
variable 9	D1995M2 - D2001M3 - D2003M1 - D2001M4 - D1996M1 - D1996M2	0.7407
variable 10	D1995M2	0.6367
variable 11		0.5547
variable 12		0.5551
variable 13	D1995M1 - D1995M2 - D1996M1 - D1996M2 - D2001M3 - D2004M2	0.8580
variable 14	D1995M1 - D1995M2 - D1996M2 - D2004M2 - D1999M4 - D2001M2	0.6688
variable 15	D1996M1 - D1996M2 - D2001M3	0.6570
variable 16	D2000M1 - D2001M3 - D2002M1	0.7935
variable 17	D2001M3 - D1995M1 - D2002M1 - D1995M2 - D2000M1 - D2001M4 - D2003M - D1996M2 - D2004M2 - D2004M2	0.8792
variable 18		0.5570
variable 19		0.5456
variable 20		0.5428
variable 21		0.6208
variable 22	D1995M1 - D1995M2 - D2001M4	0.6981
variable 23		0.5811
variable 24	D1995M1 - D1995M2 - D2001M3 - D1999M4	0.7148
variable 25	D1995M1 - D1995M2 - D2001M3	0.8013
variable 26		0.5791
variable 27		0.5904
variable 28	D1995M1 - D1995M2 - D2001M3 - D2001M4 - D1996M1 - D1996M2 - D1999M	0.8090
variable 29		0.6093
variable 30		0.5464
variable 31		0.5511
variable 32		0.5648
variable 33		0.5659
variable 34		0.6460
variable 35		0.5991
variable 36	D1995M1 - D1995M2 - D2001M3 - D2001M4 - D1999M4 - D1996M1 - D1996M	0.7776
variable 37		0.5807
variable 38		0.5653
variable 39	D1995M1 - D2001M3 - D2000M1 - D2001M4 - D2003M4	0.8629
variable 40	D1995M1 - D1995M2 - D2001M3 - D2001M4 - D1996M2	0.7566
variable 41	D1996M2 - D2003M4	0.8366
variable 42	D1995M1 - D2001M3 - D2000M1 - D1999M2	0.7935
variable 43	D1995M1 - D1995M2 - D2001M4	0.7865
variable 44	D1995M1 - D1995M2 - D2001M3 - D2001M4 - D1999M4 - D1998M3 - D1998M - D1996M2	0.8054
variable 45	D1995M1 - D1995M2 - D2001M3	0.6917
variable 46		0.6417
variable 47	D1995M1 - D1995M2 - D2001M3 - D1996M1 - D1996M2 - D2001M4 - D1998M - D1998M4 - D2003M1 - D2003M1	0.8489
variable 48	D2001M3 - D1996M1 - D1996M2	0.7457
variable 49		0.6306
variable 50		0.6288
variable 51	D1995M1 - D1995M2 - D2000M1	0.7248
variable 52		0.6881
variable 53	D1995M1 - D1995M2 - D2001M3 - D2000M1 - D2001M4	0.8187
variable 54	D2002M1 - D2003M1 - D2004M2 - D1999M4	0.6828
variable 55	D2002M2	0.75178

Note:Shaded figures indicate R^2 superior to 0.75

Selected series to construct Factors.

Table A.3.: Selected series

No. Series	Description
1	Autobile national production
9	Hot rolled nonflat steel production
13	Income Revenues Customs (DGA)
16	MERVAL stock market index monthly
17	MERVAL stock market index at month end
25	Industrial Survey - outlook non-durable cons. Goods
28	Industrial Survey - outlook intermediate goods
36	Industrial Survey - outlook durable cons. Goods
39	Industrial production index (IPI) - general level
40	IPI- non-durable consumer goods
41	IPI - durable consumer goods
42	IPI - intermediate goods
43	IPI - capital goods
44	IPI - food and beverages
47	IPI - pulp and paper
53	IPI - automobiles
55	Nominal interest rate 30-59 days - private banks