

Working Paper
Number 13/14

Short-Run Forecasting of Argentine GDP Growth

Madrid, April 2013
Economic Analysis

Short-Run Forecasting of Argentine GDP Growth*

Máximo Camacho^a, Marcos Dal Bianco^b, Jaime Martínez-Martín^c

April 2013

Abstract

We propose a small-scale dynamic factor model for monitoring Argentine GDP in real time using economic data at mixed frequencies (monthly and quarterly) which are published with different time lags. Our model not only produces a coincident index of the Argentine business cycle in striking accordance with professional consensus and the history of the Argentine business cycle, but also generates accurate short-run forecasts of the highly volatile Argentine GDP growth. By using a simulated real-time empirical evaluation, we show that our model produces reliable backcasts, nowcasts and forecasts well before the official data is released.

Keywords: Real-time forecasting, Argentine GDP, business cycles, state-space models, mixed frequencies.

JEL: C22, C53, E27, E32, E37.

*: M. Camacho would like to thank CICYT (ECO2010-19830) for their financial support. All errors are our own responsibility. The views expressed in this paper are those of the authors and do not necessarily represent those of BBVA Research. Corresponding authors: marcos.dalbianco@bbva.com and j.martinez.martin@bbva.com.

a: Universidad de Murcia.

b: BBVA Research, Argentina.

c: BBVA Research, Spain.

1 Introduction

Gross Domestic Product (GDP) is the most important measure of the aggregate state of economic activity in any market economy. As such, it should be the most relevant business cycle indicator for policymakers and economic agents who are constantly making relevant real-time economic decisions. However, the quarterly GDP figures are typically published with significant time lags, which constitutes a major problem for economic agents who need updated information in order to make a proper assessment of current and future macroeconomic conditions. Needless to say, this problem is particularly acute in emerging countries, which usually face longer publication delays for the relevant economic indicators than developed countries.

In Argentina, for instance, the GDP data for a given quarter is published about 10 weeks after the end of the corresponding quarter, clearly too late to be a useful indicator for real-time decisions. Economic agents (investors, policymakers, consumers) are therefore forced to rely on other economic series available throughout the quarter to track the evolution of current GDP. However, those series correlate with partial aspects of economic activity, are usually more volatile than GDP and often yield contradictory insights about how GDP is evolving over a given quarter. Moreover, it is difficult to use monthly series to forecast quarterly GDP since the models need to process data with different frequencies. Hence, having an econometric model that can combine monthly and quarterly economic series to obtain a real-time measure of economic activity as an updating assessment tool for tracking quarterly GDP is of utmost interest. As a result, it comes as no surprise that research economists devote increasing time and effort to developing econometric techniques to address these shortcomings.

In this context, a very useful small-scale factor model for building a coincident index of the business cycle that mixes monthly and quarterly series was initially developed by Mariano and Murasawa (2003) for the US economy. It was later refined for specific forecasting purposes by Camacho and Perez-Quiros (2010) for the eurozone. These two works, and a small but growing literature that follows them (see Section 2 below), analyze only developed regions or countries. However, to the best of our knowledge there is still no paper that applies this methodology to developing countries. This is particularly relevant as developing countries usually face much more volatile business cycles than developed ones.¹ For instance, the variance of Argentine real GDP fluctuations between 1993 and 2012 is eight times higher than that of the USA (see Figure 1).² Is it hence very interesting to see if this methodology can be successfully applied to a middle-income developing country as Argentina. In this work, we perform

¹See Lane (2003) and Aguiar and Gopinath (2007).

²However, it is important to note that in spite of its high volatility, the Argentine GDP growth does not follow an ARCH process. We carry out ARCH tests for the residual from an AR(1) specification of Argentine real GDP growth and were not able to reject the null of no-presence of ARCH in the residuals at usual significance levels. These results are available from the authors upon request.

this task. In particular, we extend those seminal works to produce backcast, nowcast and short-run forecast estimates of Argentine real GDP growth. Thus, our model uses partial information on the current economic situation mixing only several monthly and quarterly indicators with differing lengths to obtain an accurate assessment of current and future Argentine real GDP growth.

[Insert Figure 1 here]

Our main results can be summarized as follows. First, the coincident indicator has a strong performance as a business cycle indicator since it is in striking accordance with professional consensus and the history of the Argentine business cycle. Second, the percentage of the variance of the very volatile Argentine real GDP growth that is explained by the model is above 89%, indicating the high potential ability of the indicators used in the model to explain Argentine growth. Third, our pseudo-real-time analysis shows that dynamic factor models clearly outperform univariate forecasts, especially when forecasting the next unavailable figure of GDP growth. This encourages real-time forecasters to back-check the bulk of monthly real and survey data which are published in the respective quarter before the next GDP release. Against this backdrop, our model is able to produce accurate forecasts, leading us to strongly support that our model is a valid tool to be used for short-term analyses.

This paper is organized as follows. The related literature is briefly reviewed in Section 2. Section 3 presents an outline of the model, indicates how to mix frequencies, describes its dynamic properties, along with the state-space representation, and demonstrates its estimation process. Section 4 contains data description and highlights the main empirical results, both in- and out-of-sample. Finally, conclusions and proposals for future lines of research are presented in Section 5.

2 Brief review of relevant related literature

The modern literature on business cycle estimation starts with the Stock and Watson (1989, 1991) coincident index. They estimate a monthly coincident index of economic activity as the unobservable factor in a dynamic factor model for four coincident indicators: industrial production, real disposable income, hours of work and sales, with the aim of providing a formal probabilistic basis for the Burns and Mitchell (1946) coincident and leading indicators. However, the dynamic factor model advocated by these authors exhibits two important drawbacks when it is used to monitor economic activity in real time. First, their method requires balanced panels, which precluded them from using data with mixed frequency or indicators with different publication delays. Therefore, their model ignores the information contained in quarterly indicators such as real GDP, which is likely the most important business cycle indicator. Second, the index they obtain is computed as linear combinations of meaningful economic indicators; the fact that it is not related to a particular variable of interest

makes it difficult to find an economic interpretation of its level or its reactions to shocks.

To address those drawbacks, Mariano and Murasawa (2003) proposed a coincident index of business cycles with the distinctive characteristic of blending indicators published both at monthly and quarterly frequencies. They also incorporate a maximum likelihood factor analysis to the four-monthly indicators, but since their methodology is able to handle mixing frequencies, they can also include real GDP as an additional fifth coincident indicator. Moreover, they can include series of differing lengths. Their coincident index accurately captures the NBER business cycle reference dates and presents a very high statistical correlation with the Stock and Watson (1991) coincident index. Moreover, their index has an economic interpretation as the common factor component in a (latent) monthly real GDP. One drawback of Mariano and Murasawa (2003) is that they do not explore the forecasting properties of their model.

The forecasting analysis of this model is tackled later on by Camacho and Perez-Quiros (2010), who successfully modified Mariano and Murasawa's model to compute short-term forecasts of the eurozone real GDP growth in real time.³ Their small-scale dynamic factor model is able to forecast eurozone real GDP growth at least as well as (and usually better than) professional forecasters. Further developments of their work are Camacho and Perez-Quiros (2011) and Camacho and Domenech (2012) for Spain, Camacho and Martinez-Martin (2012) for the USA, and Camacho and Garcia-Serrador (2013) for the euro area.

This recent literature on short-run real GDP growth forecasting is almost exclusively focused on developed economies. The related literature is very scarce for emerging countries in general, and especially for Argentina. To the best of our knowledge, there are only three attempts similar to ours in the literature, but using different approaches. The first is Simone (2001), who constructs coincident and leading indicators of economic activity in Argentina. Although he proposes a useful contribution, he only uses quarterly data and does not attain a reliable leading indicator for Argentine real GDP. Second, two recent works by D'Amato et al. (2011a, 2011b) employ two techniques to produce predictions of current real GDP growth within the quarter, referred to as "nowcasting," and one-quarter ahead forecast of real GDP growth.⁴ Third, Liu et al. (2011) estimate a large-scale factor model based on monthly data for nowcasting and forecasting. However, the related RMSE show traces of weak forecasting capacity.

³Aruoba and Diebold (2010) also build on Stock and Watson (1989, 1991) and Mariano and Murasawa (2003) and examine the real-time performance of the common factor as a business cycle indicator, but their focus is on the assessment of current economic activity and not on forecasting.

⁴They do not present "backcasting" results, which are the estimation on a given quarter of the previous quarter rate of growth *before* they are published by the statistical agency, i.e. within the ten weeks of delay.

3 The econometric model

3.1 Mixing frequencies

We use data at two frequencies: monthly and quarterly. To mix them, we consider all series as being of monthly frequency and treat quarterly data as monthly series. In this case, the monthly series are observed in the last month of the quarter and exhibit missing observations in the first two months of each quarter.

In particular, let G_t be the level of a quarterly flow variable that can be decomposed as the sum of three (usually unobserved) monthly values G_t^* . To avoid using a non-linear state-space model, we follow Mariano and Murasawa (2003) and approximate the arithmetic mean with the geometric mean.⁵ Hence, the level of the variable can be written as

$$G_t = 3(G_t^* G_{t-1}^* G_{t-2}^*)^{1/3}. \quad (1)$$

Taking logs on both sides of this expression and computing the three-period differences for all t , we obtain

$$\Delta_3 \ln G_t = \frac{1}{3}(\Delta_3 \ln G_t^* + \Delta_3 \ln G_{t-1}^* + \Delta_3 \ln G_{t-2}^*). \quad (2)$$

Denoting the quarter-on-quarter growth rate $\Delta_3 \ln G_t = g_t$ the monthly-on-monthly growth rate $\Delta \ln G_t^* = g_t^*$ and applying algebra, we obtain

$$g_t = \frac{1}{3}g_t^* + \frac{2}{3}g_{t-1}^* + g_{t-2}^* + \frac{2}{3}g_{t-3}^* + \frac{1}{3}g_{t-4}^*. \quad (3)$$

Accordingly, we express the quarter-on-quarter growth rate (g_t) as a weighted average of the past monthly-on-monthly growth rates (g_{t-i}^* , $i = 0, \dots, 4$) of the monthly series.

3.2 Dynamic properties

We have applied the common assumption in factor modeling literature that the time series used in the model are the sum of two orthogonal components: a common component, x_t which represents the overall business cycle conditions, and an idiosyncratic component, which refers to the particular dynamics of the series. The underlying business cycle conditions are assumed to evolve with $AR(p1)$ dynamics:

$$x_t = d_1^x x_{t-1} + \dots + d_{p1}^x x_{t-p1} + \varepsilon_t^x, \quad (4)$$

where $\varepsilon_t^x = iN(0, \sigma_x^2)$.

⁵Aruoba et al. (2009) extended this analysis to include high frequency data using an exact algorithm, as opposed to the approximate algorithm of Mariano and Murasawa (2003). However, Aruoba et al. (2009) address the cost of assuming deterministic trends in the series.

For the sake of simplicity, let us assume that we consider one quarterly indicator and one monthly indicator.⁶ Starting from the quarterly indicator, let us assume that the evolution of its underlying monthly growth rates depends linearly on x_t and on the idiosyncratic dynamics, u_t^g which evolves as an $AR(p2)$:

$$g_t^* = \beta_g x_t + u_t^g, \quad (5)$$

$$u_t^g = d_1^g u_{t-1}^g + \dots + d_{p2}^g u_{t-p2}^g + \varepsilon_t^g, \quad (6)$$

where $\varepsilon_t^g = iN(0, \sigma_g^2)$. In addition, the evolution of the monthly indicator depends linearly on x_t and on the idiosyncratic component, whose dynamics can be expressed in terms of autoregressive processes of $p3$ orders:

$$z_t = \beta_z x_t + u_t^z \quad (7)$$

$$u_t^z = d_1^z u_{t-1}^z + \dots + d_{p3}^z u_{t-p3}^z + \varepsilon_t^z, \quad (8)$$

where $\varepsilon_t^z = iN(0, \sigma_z^2)$. Finally, all the shocks e_t , ε_t^y , and ε_t^z are assumed to be mutually uncorrelated in cross-section and time-series dimensions.

Using the assumptions described below, this model can be easily stated in state-space representation and estimated, as further described in the following section, using the Kalman filter.

3.3 State-space representation

Let I_a be the identity matrix of order a and let $0_{a \times b}$ be a $(a \times b)$ matrix of zeroes. For clarity purposes, let us assume that $p1 = p2 = p3 = 1$. In addition, let us start by assuming that all variables are always observed at a monthly frequency. In this simplified version, the *measurement equation*:

$$Y_t = H h_t + E_t, \quad (9)$$

with $E_t \sim i.i.d.N(0, R)$ can be stated by defining

$$Y_t = (g_t, z_t)', \quad (10)$$

$$H = \begin{pmatrix} \frac{\beta_g}{3} & \frac{2\beta_g}{3} & \beta_g & \frac{\beta_g}{3} & \frac{2\beta_g}{3} & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 \\ \beta_z & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad (11)$$

$$h_t = (x_t, x_{t-1}, x_{t-2}, x_{t-3}, x_{t-4}, u_t^g, u_{t-1}^g, u_{t-2}^g, u_{t-3}^g, u_{t-4}^g, u_t^z)', \quad (12)$$

$$E_t = (0, 0)', \quad (13)$$

$$R = 0_{2 \times 2}. \quad (14)$$

In the same way, the *transition equation*:

$$h_t = F h_{t-1} + \xi_t, \quad (15)$$

⁶Extending the model to k_1 quarterly indicators and k_2 monthly indicators is straightforward.

with $\xi_t \sim i.i.d.N(0, Q)$ can be stated by defining

$$F = \begin{pmatrix} F_1 & 0_{5 \times 5} & 0_{5 \times 1} \\ 0_{5 \times 5} & F_2 & 0_{5 \times 1} \\ 0_{1 \times 5} & 0_{1 \times 5} & d_1^z \end{pmatrix}, \quad (16)$$

$$F_1 = \begin{pmatrix} d_1^x & 0_{1 \times 4} \\ I_4 & 0_{4 \times 1} \end{pmatrix}, \quad (17)$$

$$F_2 = \begin{pmatrix} d_1^g & 0_{1 \times 4} \\ I_4 & 0_{4 \times 1} \end{pmatrix}, \quad (18)$$

$$\xi_t = (\varepsilon_t^x, 0, 0, 0, 0, \varepsilon_t^g, 0, 0, 0, 0, \varepsilon_t^z)' , \quad (19)$$

$$Q = \text{diag}(\sigma_x^2, 0, 0, 0, 0, \sigma_g^2, 0, 0, 0, 0, \sigma_z^2)' . \quad (20)$$

The identifying assumption implies that the variance of the common factor, σ_x^2 is normalized to a value of one.⁷

3.4 Estimation

The estimation of the model would be by standard maximum likelihood and using the Kalman filter if all series were observable at the monthly frequency, as we have assumed so far. However, this assumption is quite restrictive since we are using time series of different length and we are mixing monthly data with quarterly data.

Mariano and Murasawa (2003) develop a framework to easily address this issue. Following these authors, the unobserved cells can be treated as missing observations and maximum likelihood estimation of a linear Gaussian state-space model with missing observations can be applied straightforwardly after a subtle transformation of the system matrices. The missing observations can be replaced with random draws ϑ_t , whose distribution cannot depend on the parameter space that characterizes the Kalman filter.⁸ Thus, the likelihood function of the observed data and that of the data whose missing values are replaced by the random draws are equivalent up to scale. In particular, we assume that the random draws come from $N(0, \sigma_\vartheta^2)$. In addition, the measurement equation must be transformed conveniently in order to allow the Kalman filter to skip the missing observations when updating.

Let Y_{it} be the i -th element of the vector Y_t and R_{ii} be its variance. Let H_i be the i -th row of the matrix H which has ς columns and let $0_{1\varsigma}$ be a row vector of ς zeroes. The measurement equation can be replaced by the following

⁷This is a very standard assumption in factor models.

⁸Note that replacements with constants would also be valid.

expressions:

$$Y_{it}^* = \begin{cases} Y_{it} & \text{if } Y_{it} \text{ observable} \\ \vartheta_t & \text{otherwise} \end{cases}, \quad (21)$$

$$H_{it}^* = \begin{cases} H_i & \text{if } Y_{it} \text{ observable} \\ 0_{1 \times \varsigma} & \text{otherwise} \end{cases}, \quad (22)$$

$$E_{it}^* = \begin{cases} 0 & \text{if } Y_{it} \text{ observable} \\ \vartheta_t & \text{otherwise} \end{cases}, \quad (23)$$

$$R_{iit}^* = \begin{cases} 0 & \text{if } Y_{it} \text{ observable} \\ \sigma_{\vartheta}^2 & \text{otherwise} \end{cases}. \quad (24)$$

According to this transformation, the time-varying state space model can be treated as having no missing observations; therefore, the Kalman filter can be directly applied to Y_t^* , H_t^* , E_t^* , and R_t^* .

The estimation of the model's parameters can be developed by maximizing the log-likelihood of $\{Y_t^*\}_{t=1}^{t=T}$ numerically with respect to the unknown parameters in matrices. Let $\hat{\zeta}_{t|\tau}$ be the estimate of ζ_t based on information up to period τ . Let $P_{t|\tau}$ be its covariance matrix. The prediction equations are:

$$\hat{\xi}_{t|t-1} = F\hat{\xi}_{t-1|t-1}, \quad (25)$$

$$P_{t|t-1} = FP_{t-1|t-1}F' + Q. \quad (26)$$

Hence, the predicted value of Y_t with information up to $t-1$, denoted $\hat{Y}_{t|t-1}$ is:

$$\hat{Y}_{t|t-1} = H^*\hat{\xi}_{t|t-1}, \quad (27)$$

and the prediction error is:

$$\eta_{t|t-1} = Y_t^* - \hat{Y}_{t|t-1} = Y_t^* - H^*\hat{\xi}_{t|t-1}, \quad (28)$$

with covariance matrix:

$$\zeta_{t|t-1} = H^*P_{t|t-1}H^* + R_t^*. \quad (29)$$

The way missing observations are treated implies that the filter, through its implicit signal extraction process, will put no weight on missing observations in the computation of the factors.

In each iteration, the log-likelihood can be computed as

$$\log L_{t|t-1} = -\frac{1}{2} \ln \left(2\pi \left| \zeta_{t|t-1} \right| \right) - \frac{1}{2} \eta'_{t|t-1} \left(\zeta_{t|t-1} \right)^{-1} \eta_{t|t-1}. \quad (30)$$

It is worth noting that the transformed filter to handle missing observations has no impact on the model estimation. In that sense, the missing observations simply add a constant to the likelihood function of the Kalman filter process. Hence, the parameters that maximize the likelihood are achieved as if all the variables were observed.

The updating equations are:

$$\hat{\xi}_{t|t} = \hat{\xi}_{t|t-1} + P_{t|t-1} H_t^{+'} \left(\zeta_{t|t-1} \right)^{-1} \eta_{t|t-1}, \quad (31)$$

$$P_{t|t} = P_{t|t-1} - P_{t|t-1} H_t^{+'} \left(\zeta_{t|t-1} \right)^{-1} H_t^+ P_{t|t-1}. \quad (32)$$

Therefore, missing observations are from the updating recursion.

Remarkably, computing short-term forecasts in real time from this model is straightforward. The reason behind this is that one can regard the future values of the time series as a set of missing observations at the end of the sample periods. After the last observation, missing values are added to the data set and the Kalman accounts for the missing data which are replaced by forecasts. In particular, the k -period-ahead forecasts are

$$\hat{Y}_{t+k|t} = H^* \hat{\xi}_{t+k|t}, \quad (33)$$

where $\hat{\xi}_{t+k|t} = F^k \hat{\xi}_{t|t}$.

4 Empirical Results

4.1 Preliminary analysis of the data

The data employed in this paper, which was collected on September 15, 2012, spans the period from January 1993 to August 2012. Regarding the relatively wide potential set of indicators that could be used in the analysis, we only chose those that satisfy the following four properties. First, they must exhibit high statistical correlation with the real GDP growth rate, which is the target series to be estimated and predicted. Second, for a given quarter they should refer to data of that quarter, which implies that they must be published before the GDP figure becomes available in the respective quarter. Third, they must be relevant in the model from both theoretical and empirical (statistical) points of view. Finally, they must be available in at least one third of the sample.

We started the analysis with the Argentine version of the set of coincident economic indicators used in Mariano and Murasawa (2003): real quarterly GDP, monthly industrial production, quarterly employment, monthly real personal income, and real trade sales, which exhibit a strong (statistical) link with the GDP cycle. However, income was discarded because its loading factor was not statistically significant and was replaced with a Synthetic Indicator of Construction Activity, which exhibits higher correlation with GDP and a statistically significant loading factor. The potential enlargements of the data set were sequentially tested by including additional indicators such as the Consumer Confidence Index. However, the loading factors were not statistically significant and they were not included in the model.

The five indicators used in the empirical analysis and their respective publication delay are summarized in Table 1. All the variables are seasonally ad-

justed and are (weekly) stationary or transformed to be stationary.⁹ Accordingly, the quarterly indicators enter in the model in quarterly growth rates while the monthly indicators enter in monthly growth rates. Before estimating the model, the variables are standardized to have zero mean and variance equal to one. Therefore, the final forecasts are computed by multiplying the initial forecasts of the model by sample standard deviation, and then adding the sample mean.

[Insert Table 1 here]

4.2 In-sample analysis

In this section we present the results obtained by the model outlined in Section 3. In Table 2 we present the estimated values for the factor loadings which reflect the degree to which variations in each observed variable are correlated with the latent factor. As observed, all variables show statistically significant loading factors. Notably, Table 2 also shows that the percentage of variance of the actual Argentine real GDP growth that is explained by the factor is very high, reaching about 90%.

[Insert Table 2 here]

The empirical reliability of the inferred factor as an Argentine business cycle indicator is examined in Figure 2. Together with this series, the figure plots the corresponding growth rates of the Monthly Estimator of Economic Activity (EMAE from its acronym in Spanish), which is a widely accepted monthly proxy of Argentine real GDP. According to this graph the evolution of the inferred factor is in striking accord with that of EMAE.

[Insert Figure 2 here]

With the aim of meticulously checking the accuracy of the common factor as a real-time business cycle indicator, we assume that the indicator is subject to regime switches.¹⁰ For this purpose, we assume that the switching mechanism of the common factor at time t , x_t is controlled by an unobservable state variable, s_t which is allowed to follow a first-order Markov chain. In line with Hamilton (1989), a simple switching model may be specified as:

$$x_t = c_{s_t} + \sum_{j=1}^p \alpha_j x_{t-j} + \varepsilon_t, \quad (34)$$

⁹In particular, if the log of a variable appears as non-stationary according to Ng and Perron (2001) unit root tests, then the data are used in growth rates. To save space, the results are not presented but they are available from the authors upon request.

¹⁰Camacho, Perez-Quiros and Poncela (2012) show that although the fully Markov-switching dynamic factor model is generally preferred to the shortcut of computing inferences from the common factor obtained from a linear factor model, its marginal gains rapidly diminish as the quality of the indicators used in the analysis increases. This is precisely our case.

where $\varepsilon_t \sim iidN(0, \sigma^2)$.¹¹ The nonlinear behavior of the time series is governed by c_{s_t} , which is allowed to change within each of the two distinct regimes $s_t = 0$ and $s_t = 1$. The Markov-switching assumption implies that the transition probabilities are independent of the information set at $t - 1$, x_{t-1} , and of the business cycle states prior to $t - 1$. Accordingly, the probabilities of remaining in each state are:

$$p(s_t = i / s_{t-1} = j, s_{t-2} = h, \dots, x_{t-1}) = p(s_t = i / s_{t-1} = j) = p_{ij}. \quad (35)$$

Taking the maximum likelihood estimates of parameters, reported in Table 3, in the regime represented by $s_t = 0$, the intercept is positive and statistically significant while in the regime represented by $s_t = 1$ it is negative and statistically significant. Hence, we can associate the first regime with expansions and positive values of the indicator and the second regime with recessions and negative values of the indicator. According to the related literature, expansions are more persistent than downturns (estimated p_{00} and p_{11} of about 0.97 and 0.87 respectively). These estimates are in line with the well-known fact that expansions are longer than contractions, on average. Using the transition probabilities, one can derive the expected number of months that the business cycle phases prevail as $(1 - p_{ii})^{-1}$. Conditional on being in $s_t = 0$ the expected duration of a typical Argentine expansion is 33 months, and the expected duration of recession is approximately 8 months.

[Insert Table 3 here]

Finally, to check empirically whether the business cycle information that can be extracted from the common factor is in line with the historical consensus, the probabilities of being in recession (or in state $s_t = 1$) extracted from our coincident indicator along with shaded areas that refer to recessionary periods according to the reference works of Jorrat and Cerro (2000) and Jorrat (2005, 2012), are plotted in Figure 3.

[Insert Figure 3 here]

The figure shows the strong performance of the coincident indicator as a business cycle indicator for Argentina since it is in striking accord with previous estimates of turning points for the Argentine cycle. During the periods classified as recessions by this related literature, our estimated smoothed probabilities of being in recession are usually high. The only recession identified by Jorrat (2012) which is not reflected in a high probability of recession in our model is the one that appears to start in 2011:11. However, Jorrat's dating of this last recession was tentative as his methodology defines as a recession a period of five consecutive monthly falls of his Economic Activity Index and in Jorrat (2012)

¹¹Based on Camacho and Perez-Quiros (2007), we did not include any lags in the factor. We checked that the resulting model is dynamically complete in the sense that the errors are white noise.

there is only four consecutive monthly falls.¹² Hence, this recession could not be confirmed with additional information in future updates of Jorrat’s business cycle dating, which would be consistent with our results. Lastly, bearing in mind Figure 2, it shows that there is a high commonality in switch times of probabilities with Argentine business cycle phases as identified by both the common factor and EMAE.¹³

4.3 Simulated real-time analysis

Among many others, Stark and Croushore (2002) suggest that the analysis of in-sample forecasting performance of competitive models is questionable since the results can be deceptively lower when using real-time vintages. This occurs because the in-sample analysis overlooks three aspects of real-time forecasting: (i) the recursive estimation of the model parameters; (ii) the real-time data flow, i.e. the fact that data is released at different point in time; and (iii) the real-time data revisions.

However, although developing real-time data sets is conceptually simple, producing real-time vintages is, as in our case, unfeasible since the historical records of many time series are not available. In the context of dynamic factor models, an interesting alternative to the real-time forecasting analysis is the pseudo-real-time forecasting exercise suggested by, among others, Giannone, Reichlin and Small (2008). Their proposal consists of taking into account the recursive estimate of the models and the real-time data flow (and hence the publication lags) but, due to data availability constraints, it does not consider data revisions.

The proposal is based on trying to mimic as closely as possible the real-time analysis that would have been performed by a potential user of dynamic factor models when forecasting, in each period of time, on the basis of different vintages of data sets. The experiment considers that the releases of each vintage contain missing data at the end of the sample that reflects the lags in the calendar of data releases that has been summarized in Table 1. This allows us every 15 days to reproduce the typical end-of-the-sample unbalanced panel faced by the forecaster due to the lack of synchronization of the data releases. Accordingly, the experiment is labeled as "pseudo" because the vintages are not obtained in pure real-time but from the latest available data set.

The forecast performance analysis was conducted to simulate real-time forecasting. The first data vintage of this experiment refers to April 1, 2002 and, although it was collected from the information of the latest available data set, it preserved the data release calendar that a forecaster would have had on that day. Using this data vintage, we computed nine-month blocks of forecasts. Among them, some refer to the previous quarter’s GDP growth before its official release

¹²That is why the last shaded month in the graph is 2012:3 which corresponds with the last data analyzed by Jorrat (2012), but this does not imply that this apparent recession ended there as it could have lasted longer.

¹³It validates the interpretation of state $s_t = 1$ as recession and the probabilities plotted in Figure 2 as probabilities of being in recession.

(backcasts), others refer to current quarter GDP growth (nowcasts), while others refer to the next quarter’s GDP growth (forecasts). Against this backdrop, the data vintages were recursively updated on the first day and fifteenth day of each month. All parameters, factors, and so forth were then re-estimated, and nine-month blocks of backcasts, nowcasts and forecasts were then computed. The final pseudo-real-time nine-month block of forecasts was made on January 15, 2013, leading to 246 different blocks of forecasts (from first quarter 2002 to 2012).

The predictive accuracy of our model is examined in Table 4. The table shows the mean-squared forecast errors (MSE), which are the average of the deviations of the predictions from the final releases of GDP available in the data set. Results for backcasts, nowcasts and forecasts appear in the second, third and fourth columns of the table, respectively. In addition to the factor model described in Section 2, two benchmark models are included in the forecast evaluation. The former is an autoregressive model of order two (AR) which is estimated in real-time producing iterative forecasts, and the latter is a random walk (RW) model whose forecasts are equal to the average latest available real-time observations.

The MSE leads to a ranking of the competing models according to their forecasting performance. However, it is advisable to test whether the forecasts made with the dynamic factor model are significantly superior to the others models’ forecasts. To analyze whether empirical loss differences between two or more competing models are statistically significant, the last three rows of the table shows the pairwise test introduced by Diebold and Mariano (DM, 1995).

The immediate conclusion obtained when comparing the forecasts is that the dynamic factor model unequivocally outperforms the alternative forecasting models, although the magnitude of these gains depend on the forecast horizon. In the backcasting exercise, the differences between the MSE results using the factor model and the benchmark models are noticeable. The relative MSE of the dynamic factor model versus RW and AR are 0.368 to 0.350 and, according to the p -values of the DM test, the differences are statistically significant (p -values of 0.001 in both cases). The relative gains diminish with the forecasting horizon, reducing to 0.742 and 0.709 in nowcasting and to 0.880 and 0.866 in forecasting. This result is quite intuitive because the backcasts and nowcasts are computed immediately before the end of the quarter, which allow the model to use the latest available information of the respective quarter from the early available indicators. However, although the gains diminish, they are still statistically significant, according to the p -values of the DM test reported in the bottom panel of this table.

[Insert Table 4 here]

5 Conclusions

In this paper, we propose a small-scale factor model with mixed frequencies to produce accurate backcasts, nowcasts and short-run forecasts of Argentine

real GDP growth. Our model is successful not only in computing a coincident indicator, which is in striking accord with the actual history of Argentine business cycle, but also in explaining a very high percentage of the variance of actual GDP growth. Moreover, our pseudo real-time analysis shows that our dynamic factor model clearly outperforms univariate forecasts, especially when forecasting the next unavailable figure of GDP growth. This encourages real-time forecasters to back-check the bulk of monthly real and survey data which are published in the respective quarter before the next GDP release. Therefore, we strongly consider that it is a valid tool to be used to monitor the business cycle and to compute short-term forecasts of Argentine GDP growth.

References

- [1] Aguiar, M., and G. Gopinath (2007), "Emerging Market Business Cycles: The Cycle Is the Trend", *Journal of Political Economy*, Vol. 114, No. 1, Feb., pp. 69-102.
- [2] Aruoba, S.B., and F.X. Diebold (2010), "Real-Time Macroeconomic Monitoring: Real Activity, Inflation, and Interactions", *American Economic Review*, Vol. 100, No. 2, May, pp. 20-24.
- [3] Aruoba, S.B., F.X. Diebold, and C. Scotti (2009), "Real-Time Measurement of Business Conditions", *Journal of Business and Economic Statistics*, Vol. 27, No. 4, Oct., pp. 417-27.
- [4] Burns, A.F., and W.C. Mitchel (1946), "Measuring Business Cycles", NBER, New York.
- [5] Camacho, M., and R. Domenech (2012), "MICA-BBVA: A factor model of economic and financial indicators for short-term GDP forecasting", *SE-RIES: Journal of the Spanish Economic Association*, Vol. 3, pp. 475-97.
- [6] Camacho, M., and A. Garcia-Serrador (2013), "The Euro-Sting revisited: the usefulness of financial indicators to obtain euro area GDP forecasts", *Journal of Forecasting*, forthcoming.
- [7] Camacho, M., and J. Martinez-Martin (2012), "Real-time forecasting US GDP from small-scale factor models", *BBVA Working Paper* 12/10, June.
- [8] Camacho, M., and G. Perez-Quiros (2007), "Jump-and-rest effect of U.S. business cycles", *Studies in Nonlinear Dynamics and Econometrics* Vol. 11, No. 4.
- [9] Camacho, M., and G. Perez-Quiros (2010), "Introducing the Euro-Sting: short term indicator of euro area growth", *Journal of Applied Econometrics*, Vol. 25, No. 4, June/July, pp. 663-94.
- [10] Camacho, M. and G. Perez-Quiros (2011), "Spain-STING: Spain Short Term INdicator of Growth", *The Manchester School*, Vol. 79, pp. 594-616.
- [11] Camacho, M., G. Perez-Quiros, and P. Poncela (2012), "Markov-switching dynamic factor models in real-time", CEPR Discussion Papers 8866.
- [12] D'Amato, L., L. Garegnani, and E. Blanco (2011), "Nowcasting economic activity in Argentina with many predictors", *Annals of the XLVI Annual Meeting of the Argentine Association of Political Economy*, Nov.
- [13] D'Amato, L., L. Garegnani, and E. Blanco (2011), "Uso de flujos de informacion de alta frecuencia para pronosticos de corto plazo de la actividad economica en Argentina", *Investigaciones Economicas*, Vol. 64, Oct.-Dec., pp. 7-33.

- [14] Diebold F., and R. Mariano (1995), "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, Vol. 13, No. 3, Jul., pp. 253-63.
- [15] Giannone, D., L. Reichlin, and D. Small (2008), "Nowcasting: The real-time informational content of macroeconomic data", *Journal of Monetary Economics*, Vol. 55, No. 4, May, pp. 665-76.
- [16] Hamilton, J. (1989), "A new approach to the economic analysis of nonstationary time series and the business cycles" , *Econometrica*, Vol. 57, No. 2, Mar., pp. 357-84.
- [17] Jorrat, J.M., and A.M. Cerro (2000), "Computing turning point monthly probability of the Argentine economy according to the leading index, 1973-2000" , *Estudios de Economía*, Vol. 27, No. 2, pp. 279-95.
- [18] Jorrat, J.M. (2005), "Construccion de Indices Compuestos Mensuales Coincidente y Lider en Argentina", Chapter 4 in Mariana Marchionjni (Ed.), *Progresos en Econometria*, Ed. Temas and AAEP, pp. 43-100.
- [19] Jorrat, J.M. (2012), "La Economia Argentina esta en Recesion desde Noviembre de 2011", mimeo, May. A shorter non-technical version was published in Clarin, 20 May 2012.
- [20] Lane, P.R. (2003), "Business Cycles and Macroeconomic Policy in Emerging Market Economies" , *International Finance*, Vol. 6, No. 1, pp. 89-108.
- [21] Liu, P., T. Matheson, and R. Romeu (2011), "Real-Time Forecasts of Economic Activity for Latin American Economies", IMF Working Paper N° 11/98.
- [22] Mariano, R.S., and Y. Murasawa (2003), "A new coincident index of business cycles based on monthly and quarterly series", *Journal of Applied Econometrics*, Vol. 18, No. 4, Jul.-Aug., pp. 427-43.
- [23] Ng, S., and P. Perron (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, Vol. 69, No. 6, Nov., pp. 1519-54.
- [24] Simone, A. (2001), "In Search of Coincident and Leading Indicators of Economic Activity in Argentina", IMF Working Paper WP/03/30, March.
- [25] Stark, T., and D. Croushore (2002), "Forecasting with Real-Time Data Set for Macroeconomists", *Journal of Macroeconomics*, Vol. 24, pp. 507-31.
- [26] Stock, J.H., and M. Watson (1989), "New Indexes of Coincident and Leading Economic Indicators", in O. Blanchard and S. Fischer (Eds.), *NBER Macroeconomics Annual*, Vol. 4, pp. 351-409.

- [27] Stock, J.H., and M. Watson (1991), "A Probability Model of the Coincident Economic Indicators", in K. Lahiri and G. Moore (Eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*, Cambridge University Press, UK, pp. 63-89.

Table 1: Final variables included in the model

	Series	Sample	Source	Publication delay	Data transformation
1	Real Gross Domestic Product (GDP, SAAR, Mil.1993 Pesos)	1993.1 2012.3	INDEC	2.5 to 3 months	QGR
2	Industrial Production Index (IPI) (SA, 1993=100)	1993.01 2012.07	FIEL	25 days	MGR
3	All Employees: Total Urban Population (Empl, SA, Thous)	1993.1 2012.2	INDEC	1.5 to 2 months	QGR
4	Real Retail Sales: Total Supermarket Sales (Sales, SA, IPC deflated, constant ARS)	1997.06 2012.06	INDEC/ Census	25-30 days	MGR
5	Synthetic Indicator of Construction Activity (ISAC, SA)	1993.01 2012.07	INDEC	30 days	MGR

Notes: SA means seasonally adjusted. MGR and QGR mean monthly growth rates, quarterly growth rates and levels, respectively. INDEC: National Institute of Statistics and Census; FIEL: Latin American Foundation of Economic Investigations.

Table 2: Loading factors

GDP	IP	Empl	ISAC	Sales	% Var
0.28 (0.05)	0.41 (0.09)	0.28 (0.12)	0.35 (0.07)	0.16 (0.08)	89.9%

Notes: The loading factors (standard errors are in brackets) measure the correlation between the common factor and each of the indicators appearing in columns. See Table 1 for a description of these indicators.

Table 3: Markov-switching estimates

c_0	c_1	σ^2	p_{00}	P_{11}
1.09 (0.15)	-5.39 (0.43)	3.45 (0.32)	0.97 (0.01)	0.87 (0.06)

Notes. The estimated model is $x_t = c_{s_t} + \varepsilon_t$, where x_t is the common factor, s_t is an unobservable state variable that governs the business cycle dynamics, $\varepsilon_t \sim iidN(0, \sigma)$, and $p(s_t = i / s_{t-1} = j) = p_{ij}$.

Table 4: Predictive accuracy

	Backcasts	Nowcasts	Forecasts
Mean Squared Errors			
Our model	1.049	1.552	1.884
RW	2.851	2.090	2.139
Our model/RW	0.368	0.742	0.880
AR	2.999	2.189	2.174
Our model/AR	0.350	0.709	0.866
Equal predictive accuracy tests			
Our model vs RW	0.001	0.032	0.041
Our model vs AR	0.001	0.015	0.021

Notes: The forecasting sample is 2002.1-2012.1, which implies comparisons over 246 forecasts. The top panel shows the Mean Squared Errors (MSE) of our dynamic factor model, a random walk (RW), an autoregressive model of order two (AR), along with the relative MSEs over that of our model. The bottom panel shows the p -values of the Diebold-Mariano (DM) test of equal predictive accuracy.

Figure 1. Argentina and USA, real GDP growth (var. % q/q, S.A., 1993:1-2012:4)

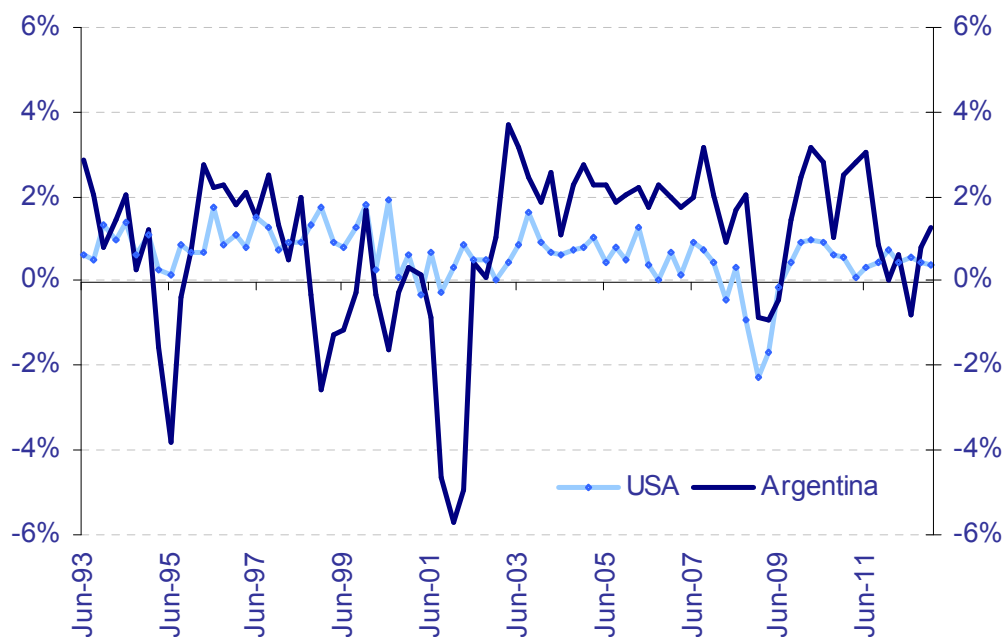
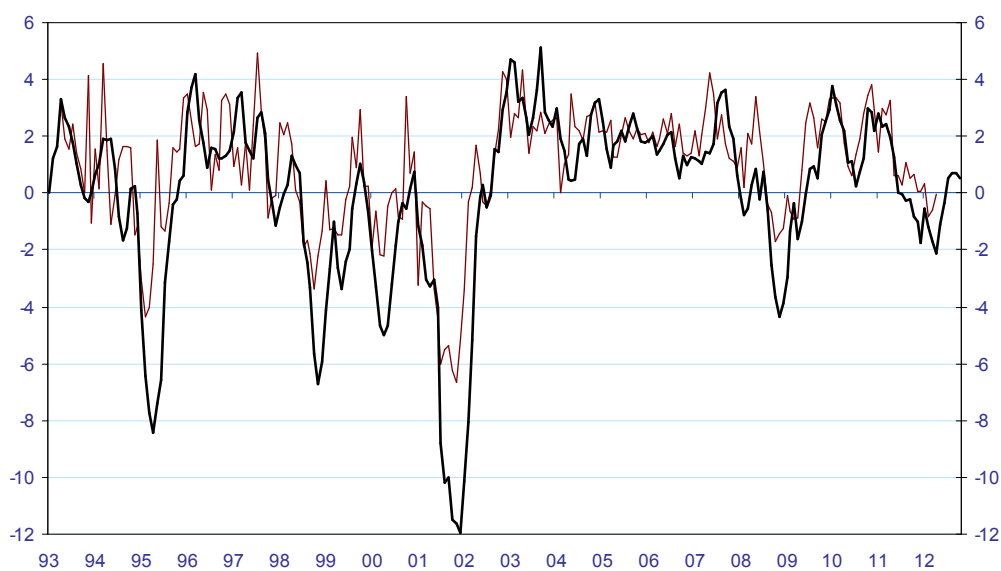
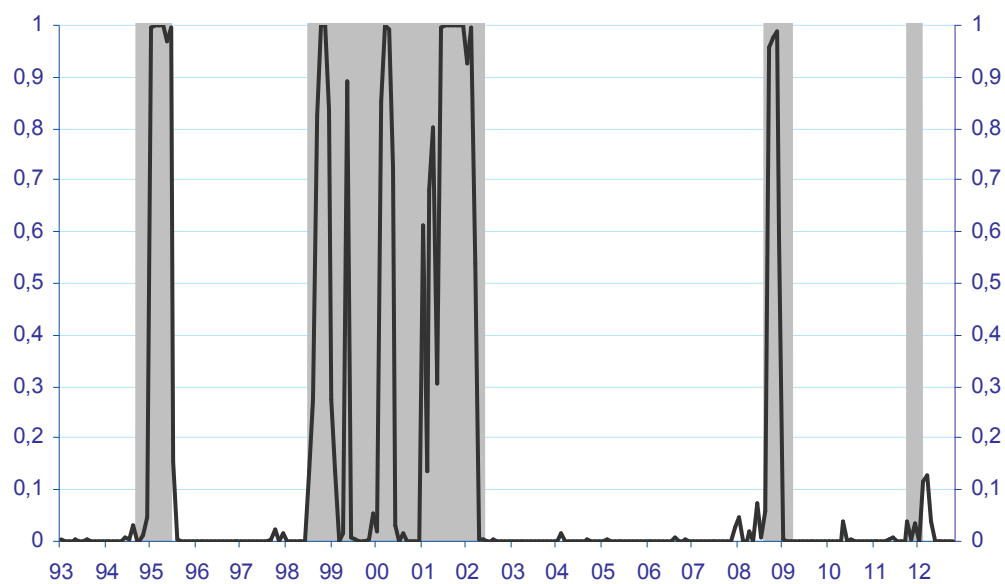


Figure 2. Common factor and EMAE (var. % 3M-3M)



Notes. Black line refers to the common dynamic factor (1993.03-2012.12). Red line refers to EMAE (1993.06-2012.06).

Figure 3. Probability of recession from our factor and periods of recession (shaded) according to Cerro and Jorrat (2000), and Jorrat (2005,12)



Working Papers

2013

13/01 **Hugo Perea, David Tuesta y Alfonso Ugarte:** Lineamientos para impulsar el Crédito y el Ahorro. Perú.

13/02 **Ángel de la Fuente:** A mixed splicing procedure for economic time series.

13/03 **Ángel de la Fuente:** El sistema de financiación regional: la liquidación de 2010 y algunas reflexiones sobre la reciente reforma.

13/04 **Santiago Fernández de Lis, Adriana Haring, Gloria Sorensen, David Tuesta, Alfonso Ugarte:** Lineamientos para impulsar el proceso de profundización bancaria en Uruguay.

13/05 **Matt Ferchen, Alicia Garcia-Herrero and Mario Nigrinis:** Evaluating Latin America's Commodity Dependence on China.

13/06 **K.C. Fung, Alicia Garcia-Herrero, Mario Nigrinis Ospina:** Latin American Commodity Export Concentration: Is There a China Effect?.

13/07 **Hugo Perea, David Tuesta and Alfonso Ugarte:** Credit and Savings in Peru.

13/08 **Santiago Fernández de Lis, Adriana Haring, Gloria Sorensen, David Tuesta, Alfonso Ugarte:** Banking penetration in Uruguay.

13/09 **Javier Alonso, María Lamuedra y David Tuesta:** Potencialidad del desarrollo de hipotecas inversas: el caso de Chile.

13/10 **Ángel de la Fuente:** La evolución de la financiación de las comunidades autónomas de régimen común, 2002-2010.

13/11 **Javier Alonso, María Lamuedra y David Tuesta:** Potentiality of reverse mortgages to supplement pension: the case of Chile.

13/12 **Javier Alonso, David Tuesta, Diego Torres y Begoña Villamide:** Proyecciones de tablas generacionales dinámicas y riesgo de longevidad en Chile.

13/13 **Alicia Garcia Herrero and Fielding Chen:** Deleveraging from Emerging Markets: the case of Euro-area Banks.

13/14 **Maximo Camacho, Marcos Dal Bianco, Jaime Martínez-Martín:** Short-Run Forecasting of Argentine GDP Growth.

2012

12/01 **Marcos Dal Bianco, Máximo Camacho and Gabriel Pérez-Quiros:** Short-run forecasting of the euro-dollar exchange rate with economic fundamentals.

12/02 **Guoying Deng, Zhigang Li and Guangliang Ye:** Mortgage Rate and the Choice of Mortgage Length: Quasi-experimental Evidence from Chinese Transaction-level Data.

12/03 **George Chouliarakis and Mónica Correa-López:** A Fair Wage Model of Unemployment with Inertia in Fairness Perceptions.

12/04 **Nathalie Aminian, K.C. Fung, Alicia García-Herrero, Francis NG:** Trade in services: East Asian and Latin American Experiences.

12/05 **Javier Alonso, Miguel Angel Caballero, Li Hui, María Claudia Llanes, David Tuesta, Yuwei Hu and Yun Cao:** Potential outcomes of private pension developments in China (Chinese Version).

12/06 **Alicia Garcia-Herrero, Yingyi Tsai and Xia Le:** RMB Internationalization: What is in for Taiwan?

- 12/07 **K.C. Fung, Alicia García-Herrero, Mario Nigrinis Ospina:** Latin American Commodity Export Concentration: Is There a China Effect?
- 12/08 **Matt Ferchen, Alicia García-Herrero and Mario Nigrinis:** Evaluating Latin America's Commodity Dependence on China.
- 12/09 **Zhigang Li, Xiaohua Yu, Yinchu Zeng and Rainer Holst:** Estimating transport costs and trade barriers in China: Direct evidence from Chinese agricultural traders.
- 12/10 **Maximo Camacho and Jaime Martinez-Martin:** Real-time forecasting US GDP from small-scale factor models.
- 12/11 **J.E. Boscá, R. Doménech and J. Ferri:** Fiscal Devaluations in EMU.
- 12/12 **Ángel de la Fuente and Rafael Doménech:** The financial impact of Spanish pension reform: A quick estimate.
- 12/13 **Biliana Alexandrova-Kabadjova, Sara G. Castellanos Pascacio, Alma L. García-Almanza:** The Adoption Process of Payment Cards -An Agent- Based Approach:
- 12/14 **Biliana Alexandrova-Kabadjova, Sara G. Castellanos Pascacio, Alma L. García-Almanza:** El proceso de adopción de tarjetas de pago: un enfoque basado en agentes.
- 12/15 **Sara G. Castellanos, F. Javier Morales y Mariana A. Torán:** Análisis del Uso de Servicios Financieros por Parte de las Empresas en México: ¿Qué nos dice el Censo Económico 2009?
- 12/16 **Sara G. Castellanos, F. Javier Morales y Mariana A. Torán:** Analysis of the Use of Financial Services by Companies in Mexico: What does the 2009 Economic Census tell us?
- 12/17 **R. Doménech:** Las Perspectivas de la Economía Española en 2012.
- 12/18 **Chen Shiyuan, Zhou Yinggang:** Revelation of the bond market (Chinese version).
- 12/19 **Zhouying Gang, Chen Shiyuan:** On the development strategy of the government bond market in China (Chinese version).
- 12/20 **Ángel de la Fuente and Rafael Doménech:** Educational Attainment in the OECD, 1960-2010.
- 12/21 **Ángel de la Fuente:** Series enlazadas de los principales agregados nacionales de la EPA, 1964-2009.
- 12/22 **Santiago Fernández de Lis and Alicia García-Herrero:** Dynamic provisioning: a buffer rather than a countercyclical tool?.
- 12/23 **Ángel de la Fuente:** El nuevo sistema de financiación de las Comunidades Autónomas de régimen común: un análisis crítico y datos homogéneos para 2009 y 2010.
- 12/24 **Beatriz Irene Balmaseda Pérez y Lizbeth Necoechea Hasfield:** Metodología de estimación del número de clientes del Sistema Bancario en México.
- 12/25 **Ángel de la Fuente:** Series enlazadas de empleo y VAB para España, 1955-2010.
- 12/26 **Oscar Arce, José Manuel Campa y Ángel Gavilán:** Macroeconomic Adjustment under Loose Financing Conditions in the Construction Sector.
- 12/27 **Ángel de la Fuente:** Algunas propuestas para la reforma del sistema de financiación de las comunidades autónomas de régimen común.
- 12/28 **Amparo Castelló-Climent, Rafael Doménech:** Human Capital and Income Inequality: Some Facts and Some Puzzles.

12/29 **Mónica Correa-López y Rafael Doménech:** La Internacionalización de las Empresas Españolas.

12/30 **Mónica Correa-López y Rafael Doménech:** The Internationalisation of Spanish Firms.

12/31 **Robert Holzmann, Richard Hinz and David Tuesta:** Early Lessons from Country Experience with Matching Contribution Schemes for Pensions.

12/32 **Luis Carranza, Ángel Melguizo and David Tuesta:** Matching Contributions for Pensions in Colombia, Mexico, and Peru: Experiences and Prospects.

12/33 **Robert Holzmann, Richard Hinz y David Tuesta:** Primeras lecciones de la experiencia de países con sistemas de pensiones basados en cotizaciones compartidas.

12/34 **Luis Carranza, Ángel Melguizo y David Tuesta:** Aportaciones compartidas para pensiones en Colombia, México y Perú: Experiencias y perspectivas.

[Click here to access the list of Working Papers published between 2009 and 2011](#)

The analysis, opinions, and conclusions included in this document are the property of the author of the report and are not necessarily property of the BBVA Group

BBVA Research's publications can be viewed on the following website: <http://www.bbvaresearch.com>

Contact details

BBVA Research
Paseo Castellana, 81 - 7th floor
28046 Madrid (Spain)
Tel.: +34 91 374 60 00 and +34 91 537 70 00
Fax: +34 91 374 30 25
bbvaresearch@grupobbva.com
www.bbvaresearch.com