

**REAL-TIME FORECASTING
US GDP FROM SMALL-SCALE
FACTOR MODELS**

2014

Máximo Camacho and Jaime Martínez-Martín

**Documentos de Trabajo
N.º 1425**

BANCO DE ESPAÑA
Eurosistema



REAL-TIME FORECASTING US GDP FROM SMALL-SCALE FACTOR MODELS

REAL-TIME FORECASTING US GDP FROM SMALL-SCALE FACTOR MODELS ^(*)

Máximo Camacho ^(**)

UNIVERSIDAD DE MURCIA

Jaime Martínez-Martín ^(***)

BANCO DE ESPAÑA

(*) We would like to thank R. Domenech, N. Karp, H. Danis, the editor and two anonymous referees for their helpful comments. We are grateful to CICYT (ECO2010-19830) and BBVA for their financial support. Both authors were working for BBVA Research during the writing of this paper. All the remaining errors are our own responsibility.

(**) Máximo Camacho, Universidad de Murcia, Facultad de Economía y Empresa, Departamento de Métodos Cuantitativos para la Economía, 30100, Murcia, Spain. E-mail: mcamacho@um.es

(***) Jaime Martínez-Martín, Banco de España, Dirección General Adjunta de Asuntos Internacionales. E-mail: jaimemartinezm@bde.es

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: <http://www.bde.es>.

Reproduction for educational and non-commercial purposes is permitted provided that the source is acknowledged.

© BANCO DE ESPAÑA, Madrid, 2014

ISSN: 1579-8666 (on line)

Abstract

We show that the single-index dynamic factor model developed by Aruoba and Diebold (Am Econ Rev, 100:20-24, 2010) to construct an index of US business cycle conditions is also very useful for forecasting US GDP growth in real time. In addition, we adapt the model to include survey data and financial indicators. We find that our extension is unequivocally the preferred alternative for computing backcasts. In nowcasting and forecasting, our model is able to forecast growth as well as AD and better than several baseline alternatives. Finally, we show that our extension could also be used to infer US business cycles with great accuracy.

Keywords: real-time forecasting, economic indicators, business cycles.

JEL Classification: E32, C22, E27.

Resumen

En este trabajo mostramos que el modelo de factores dinámicos de un solo índice desarrollado por Aruoba y Diebold (*American Economic Review*, 100, pp. 20-24, 2010) para construir un indicador de condiciones del ciclo económico de Estados Unidos resulta también muy útil para predecir el crecimiento del PIB de Estados Unidos en tiempo real. Además, adaptamos el modelo para incluir datos de encuestas e indicadores financieros. Encontramos que nuestra extensión es inequívocamente la alternativa preferida para calcular *backcasts*. En *nowcasting* y *forecasting*, nuestro modelo es capaz de predecir el crecimiento tan bien como Aruoba y Diebold y mejor que varias alternativas de base. Finalmente, mostramos que nuestra extensión podría ser también utilizada para inferir los ciclos económicos de Estados Unidos de forma muy precisa.

Palabras clave: predicción en tiempo real, indicadores económicos, ciclos económicos.

Códigos JEL: E32, C22, E27.

1 Introduction

The Great Recession of 2008/9 came as a big shock to policy makers and the business world. The rapid downturn in the economy triggered drastic reactions by policy makers who implemented monetary and fiscal policies to combat the adverse economic situation. In addition, the pervasive effects on retirement plans, stock portfolios and part-time work drastically changed private agents' economic decisions. Since being late entailed dramatic economic consequences, the economic agents seemed to learn the lesson when the recovery started. They acknowledged the need for new tools to monitor economic developments in real time.

In the context of the US economy, Auroba and Diebold (AD, 2010) is an excellent contribution to the warming debate. In line with the seminal proposal of Stock and Watson (1991), they use a small-scale single-index dynamic factor model to produce an accurate economic indicator of US business conditions in real time. Apart from the quarterly GDP growth, their model benefits from the information provided by four monthly coincident economic indicators, industrial production, payroll employment, real personal income less transfers, and trade sales.¹ Using the method proposed by Mariano and Murasawa (2003), AD adjust the factor model to handle the mixing frequencies problem and the different start and finish dates of the indicators, as they are typical in real-time forecasting due to differing release timeliness.

Although AD find that the movements in the real activity indicator strongly cohere with the NBER chronology, plunging during recessions and recovering its average level during expansions, they did not exploit the potential usefulness of the dynamic factor model to forecast US GDP growth rates in real time. The first purpose of this paper is to fill this gap by computing short-term forecasts from the model which are compared with several standard forecasting alternatives. The analysis is developed thorough a pseudo real-time analysis where the data vintages are constructed by taking into account the lag of synchronicity in data publication that characterizes real-time data flow. In addition, according to the standard literature on forecasting, the forecasts are computed in a recursive way. Therefore, although the vintages are constructed from the latest available data, the models are re-estimated and the forecasts for different horizons are computed with every new vintage that accounts for different data releases as it would have been done by a real-time forecaster.

In addition, a second purpose of the paper is to examine whether it is worth enlarging this basic dynamic factor model by taking into account the potential forecasting improvements of soft and financial indicators, which have the appealing of exhibiting very short publication delays. Since financial indicators could lead real activity, to examine this question the baseline model is extended to include leading along with coincident indicators, following the lines suggested by Camacho and Domenech (2012).

Our main results can be summarized as follows. First, we ascertain that the percentage of the variance of GDP growth that is explained by AD is slightly above 75%, indicating the high potential ability of the dynamic factor model to explain US growth. Second,

¹. AD is based on Aruoba, Diebold and Scotti (2009). Their latest version adds a weekly indicator, which implies that the model is computationally demanding to handle the high-frequency indicator. Since we performed thousands of pseudo real-time forecasts, we focus on AD which only uses monthly and quarterly indicators.

our pseudo real-time analysis shows that this dynamic factor model clearly outperforms univariate forecasts, especially when forecasting the next unavailable figure of GDP growth. Third, our extension of the dynamic factor model produces the most accurate backcasts, and leads to as accurate nowcasts and forecasts as AD. 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. Fourth, we find a high performance of the enlarged coincident indicator as a business cycle indicator since it (as well as AD) is in striking accord with the professional consensus of the history of the US business cycle.

The structure of this paper is as follows. Section 2 outlines the dynamic factor model, shows how to mix frequencies, states the time series dynamic properties, and describes the state space representation. Section 3 contains data description and the main empirical results. Section 4 concludes and proposes several future lines of research.

2 The model

2.1 Mixing frequencies

Let us assume that the level of GDP in quarter τ , Y_{τ}^* , can be decomposed as the sum of three unobservable monthly values Y_t, Y_{t-1}, Y_{t-2} , where $t, t-1$ and $t-2$ refer to the three months of quarter τ in this case. For instance, the GDP for the third quarter of a given year is the sum of the GDP corresponding to the three months of the third quarter

$$Y_{III}^* = Y_{09} + Y_{08} + Y_{07}, \quad (1)$$

or equivalently

$$Y_{III}^* = 3 \left(\frac{Y_{09} + Y_{08} + Y_{07}}{3} \right). \quad (2)$$

Among others, Mariano and Murasawa (2003) have shown that if the sample mean of equation (2) can be well approximated by the geometric mean

$$Y_{III}^* = 3(Y_{09}Y_{08}Y_{07})^{1/3}, \quad (3)$$

then the quarterly growth rates can be decomposed as weighted averages of monthly growth rates. Taking logs of expression (3) leads to

$$\ln Y_{III}^* = \ln 3 + \frac{1}{3}(\ln Y_{09} + \ln Y_{08} + \ln Y_{07}), \quad (4)$$

which allows us to compute the quarterly growth rate for the third quarter as

$$\begin{aligned} \ln Y_{III}^* - \ln Y_{II}^* &= \frac{1}{3}(\ln Y_{09} + \ln Y_{08} + \ln Y_{07}) - \frac{1}{3}(\ln Y_{06} + \ln Y_{05} + \ln Y_{04}) = \\ &= \frac{1}{3}[(\ln Y_{09} - \ln Y_{06}) + (\ln Y_{08} - \ln Y_{05}) + (\ln Y_{07} - \ln Y_{04})], \end{aligned} \quad (5)$$

and by redefining these terms as $y_{III}^* = \ln Y_{III}^* - \ln Y_{II}^*$, and $y_j = \ln Y_j - \ln Y_{j-1}$, one can define

$$y_{III}^* = \frac{1}{3}y_{09} + \frac{2}{3}y_{08} + y_{07} + \frac{2}{3}y_{06} + \frac{1}{3}y_{05}. \quad (6)$$

Calling y_{τ}^* the quarter-over-quarter growth rate in quarter τ , and y_t the respective month-over-month growth rate that refers to the last month of the quarter, this expression can directly be generalized as

$$y_{\tau}^* = \frac{1}{3}y_t + \frac{2}{3}y_{t-1} + y_{t-2} + \frac{2}{3}y_{t-3} + \frac{1}{3}y_{t-4} \quad (7)$$

This aggregation rule represents the quarterly growth rate as the weighted sum of five monthly growth rates.

2.2 Dynamic properties

The model follows the lines proposed by Camacho and Perez Quiros (2010) and Aruoba and Diebold (2010), which are extensions of the dynamic factor model suggested by Stock and Watson (1991). Let us assume that the indicators included in the model admit a dynamic factor representation. In this case, the variables can be written as the sum of two stochastic 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

$$x_t = \rho_1 x_{t-1} + \dots + \rho_{p1} x_{t-p1} + e_t, \quad (8)$$

where
$$e_t \sim iN(0, \sigma_e^2).$$

Apart from constructing an index of the business cycle conditions, we are interested in computing accurate short-term forecasts of GDP growth rates. To compute these forecasts, we start by assuming that the evolution of the monthly growth rates depend linearly on x_t and on their idiosyncratic dynamics, u_t^y , which evolve as an $AR(p2)$

$$y_t = \beta_y x_t + u_t^y, \quad (9)$$

$$u_t^y = d_1^y u_{t-1}^y + \dots + d_{p2}^y u_{t-p2}^y + \varepsilon_t^y, \quad (10)$$

where $\varepsilon_t^y \sim iN(0, \sigma_y^2)$. In addition, the idiosyncratic dynamics of the k monthly indicators can be expressed in terms of autoregressive processes of $p3$ orders:

$$z_t^i = \beta_i x_t + u_t^i, \quad (11)$$

$$u_t^i = d_1^i u_{t-1}^i + \dots + d_{p3}^i u_{t-p3}^i + \varepsilon_t^i, \quad (12)$$

where $\varepsilon_t^i \sim iN(0, \sigma_i^2)$. Finally, we assume that all the shocks e_t , ε_t^y , and ε_t^i , are mutually uncorrelated in cross-section and time-series dimensions.

2.3 State space representation

Let us first assume that all the variables included in the model were observed at monthly frequencies for all periods. Let us show how the state space representation handles the mixing frequencies problem by bridging monthly and quarterly data.² Since GDP is used in quarterly growth rates, y_t^* , according to expressions (7)-(9) it enters into the model as

$$y_t^* = \beta_y \left(\frac{1}{3} x_t + \frac{2}{3} x_{t-1} + x_{t-2} + \frac{2}{3} x_{t-3} + \frac{1}{3} x_{t-4} \right) + \left(\frac{1}{3} u_t^y + \frac{2}{3} u_{t-1}^y + u_{t-2}^y + \frac{2}{3} u_{t-3}^y + \frac{1}{3} u_{t-4}^y \right). \quad (13)$$

The unit roots of hard indicators are accounted for by using the time series in their monthly growth rates. Soft indicators, such as the consumer confidence and the purchasing

² It is worth noting that AD used the Harvey aggregator. In our application, we checked that the empirical differences between these two aggregation methods are negligible.

managers' index, are used in levels. Calling Z_t^* the monthly growth rates of hard or the level of soft variables, the dynamics of these variables are captured by

$$Z_{it}^* = \beta_i x_t + u_t^i, \quad (14)$$

with $i = 1, 2, \dots, k1$.

Finally, following the suggestions of Wheelock and Wohar (2009), financial indicators are treated as leading indicators of the current business conditions.³ Accordingly, following the lines suggested by Camacho and Domenech (2012), we establish the relationship between the level (in the case of term spread) of the financial indicator, Z_{ft}^* , and the h -period future values of the common factor, as follows:

$$Z_{ft}^* = \beta_f x_{t+h} + u_t^f. \quad (15)$$

As it is shown in the Appendix, this model can be easily stated in state space representation and estimated by using the Kalman filter. However, we assumed that the time series do not contain missing data which becomes clearly an unrealistic assumption since our data exhibits ragged ends and mixing frequency problems. Fortunately, Mariano and Murasawa (2003) show that the Kalman filter can be used to estimate the model's parameters and infer unobserved components and missing observations. These authors propose replacing the missing observations with random draws \mathcal{G}_t , whose distribution cannot depend on the parameter space that characterizes the Kalman filter.⁴ Hence, while this procedure leaves the matrices used in the Kalman filter conformable, the rows containing missing observations will be skipped from the updating in the recursions and the missing data are replaced by estimates. In this way, forecasting is very simple since forecasts can be viewed as missing data located at the end of the model's indicators.

3. To facilitate the analysis, following Giannone, Reichlin and Small (2008) financial data are entered into the model as monthly averages since the bulk of information compiled from the indicators is monthly.

4. We assume that $\mathcal{G}_t \sim N(0, \sigma_{\mathcal{G}}^2)$ for convenience but replacements by constants would also be valid.

3 Empirical results

3.1 Preliminary analysis of data

The data set managed in this paper, which was collected on January 29, 2012, spans the period from January 1960 to December 2011. The indicators used in the empirical analysis and their respective release lag-time are listed in Table 1. All the variables are seasonally adjusted. GDP enters in the model as its quarterly growth rate; hard indicators enter in monthly growth rates; and soft and financial indicators enter with no transformation. Before estimating the model, the variables are standardized to have a zero mean and a variance equal to one. Therefore, the final forecasts are computed by multiplying the initial forecasts of the model by the sample standard deviation, and then adding the sample mean.⁵ In the simulated real-time analysis, the sample means and standard deviations are also computed using only the observations available up to the forecast jump-off point.

Table 1: Final variables included in the model

	Series	Sample	Source	Publication delay	Data transform.
1	Real Gross Domestic Product (GDP, SAAR, Bil.Chn.2005\$)	60.1 11.4	BEA	3	QGR
2	Industrial Production Index (IPI) (SA, 2007=100)	60.01 11.12	Fed Reserve	2	MGR
3	All Employees: Total Nonfarm Payrolls (Empl, SA, Thous)	60.01 11.12	BLS	1.5	MGR
4	Real Personal Income Less Transfer Payments (Income, SAAR, Bil.Chn.2005\$)	60.01 11.11	BEA	2	MGR
5	Real Manufacturing Trade and Trade Sales (Sales, SA, Mil.\$)	67.01 11.10	BEA	2	MGR
6	Mfrs' New Orders: Nondefense Capital Goods ex Aircraft (MNO, SA, Mil.\$)	92.03 11.12	Census	0	MGR
7	Conference Board: Consumer Confidence (CC, SA, 1985=100)	67.02 11.12	Conference Board	0	L
8	ISM Manufacturing: PMI Composite Index (PMI, SA, 50+=Increasing)	60.01 11.12	ISM	0	L
9	House Housing Starts (House, SAAR, Thous.Units)	60.01 11.12	Census	2	MGR
10	Standard & Poor's 500 Stock Price Index (SP500, 1941-43=10)	60.01 11.12	NYT	0	MGR
11	Slope Yield Curve 10Y-Fed (Slope)	62.01 11.12	Treasury & FRB	0	L

Notes. SA means seasonally adjusted. MGR, QGR and L mean monthly growth rates, quarterly growth rates and levels, respectively.

5. To simplify the analysis, all the dynamic factor models use $p_1 = p_2 = p_3 = 2$.

3.2 In-sample analysis

Selecting the indicators that must be included in a dynamic factor model from the universe of potentially available time series is still an open question in empirical studies. For instance, Boivin and Ng (2006), have found that selecting a smaller subset of the potential set of available indicators, and using the factors that summarize the information in that smaller subset of data in the forecasting equation, substantially improves forecast performance.

Towards this end, we started the analysis with the set of coincident economic indicators used in AD, real quarterly GDP, monthly industrial production, payroll employment, real personal income less transfers, and trade sales, which exhibit a strong link with the GDP cycle. The estimated loading factors of this model are displayed in the row labeled as M2 in Table 2. Notably, the loading factors of the monthly indicators are quite similar to those displayed in row M1, which correspond to the model that does not use GDP as in Stock and Watson (1991). All of them are positive and statistically significant. The loading factor of real GDP is also positive and statistically significant. Therefore, these economic indicators are procyclical. The percentage of the variance of GDP that is explained by the model stands slightly above 75%, indicating the high potential ability of the indicators used in the model to explain GDP.

Regarding the potential set of indicators that could be used to extend the AD model, we only consider those that verify four properties. First, they must exhibit high statistical correlation with the GDP growth rate, which is the target series to be predicted. Second, for a given quarter they should refer to data of this quarter, which must be published before the GDP figure becomes available in the respective quarter. Third, they must be available in at least one third of the sample. Finally, they must be relevant in the model from both theoretical and empirical points of view.

To decide if an indicator is finally included in the model, we follow the recommendations suggested by Camacho and Perez Quiros (2010). They propose that a candidate is further added to the estimation whenever (i) it does not reduce significantly the percentage of the variance of GDP explained by the common factor; and (ii) it exhibits a statistically significant loading factor. Otherwise, the information provided by the potential indicator is assumed to be mainly idiosyncratic and it is not included in the model.

In this search, we focus our attention on two types of business cycle indicators. The first set of indicators comprises those indicators that exhibit short publication delays. Among them, we include early published hard (economic activity) indicators, which are typically available with a delay of one or two months, and soft (based on opinion surveys) indicators, which do not exhibit publication delays. Among the set of hard indicators we include new industrial orders and housing starts. Among the set of soft indicator, we include the Conference Board consumer confidence index and the ISM manufacturing PMI. According to the rows labeled as M3, to M5 in Table 2, the loading factors of these indicators are positive and statistically significant and the percentage of GDP explained by the model increases to 77.14 in M5.

The second set of indicators refers to financial indicators, which are available on a timely basis. In particular, we focus on SP500 and the term spread, which is measured as the difference between the yields on long-term and short-term maturities (10-year Treasury bond yield at constant maturity minus Federal Funds effective rate). As we discussed in Section 2, we examine the extent to which these financial indicators lead the business cycle dynamics in

h months, with $h=0, 3, 6, 9$ and 12 .⁶ To select the optimal number of leads, we compute the log likelihood values associated with these lead times, which are plotted in Figure 1.

The left-hand-side chart shows that the maximum of the likelihood function is achieved when SP500 enters as a coincident indicator of the common factor. Accordingly, model M6 of Table 2 displays the loading factors and the percentage of the variance of GDP that is explained by the common factor by using $h=0$. The table shows that SP500 exhibits positive and statistically significant loading factor and it increases the variance of GDP explained by the factor up to 78.12%. Therefore, SP500 is included among the set of indicators.

Regarding the term spread, the right-hand-side of Figure 1 shows that the maximum of the likelihood function is achieved when the spread leads the common factor by $h=3$ months. The estimated loading factor of the model that includes the term spread leading the factor by three months, which is displayed in the row labeled as M7 in Table 2, shows that it is not statistically significant. Therefore, the term spread is not included in the model and model M6 is our selected model.⁷

Table 2: Loading factors

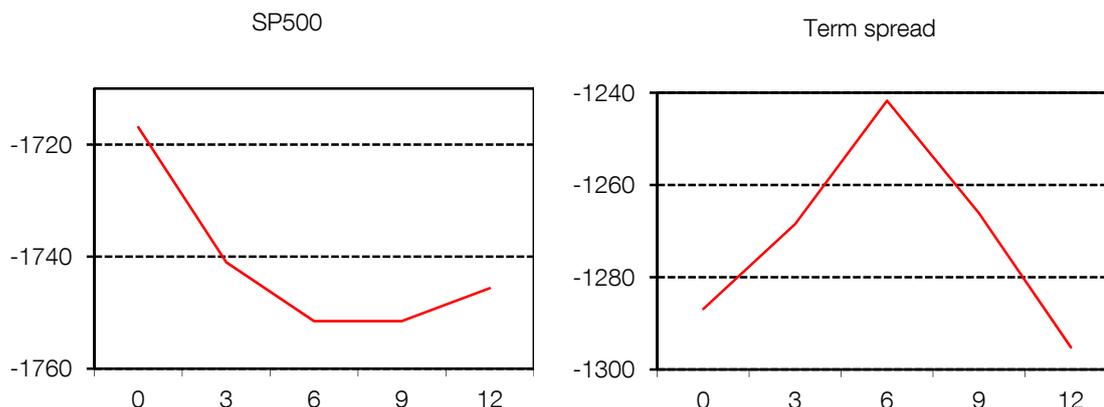
GDP	IP	Empl	Inc	Sales	MNO	CC	PMI	House	SP500	Slope	% var
---	0.62 (0.03)	0.55 (0.03)	0.35 (0.03)	0.44 (0.03)	---	---	---	---	--	---	---
0.25 (0.01)	0.59 (0.03)	0.56 (0.03)	0.37 (0.03)	0.21 (0.02)	---	---	---	---	--	---	76.64%
0.25 (0.01)	0.60 (0.03)	0.55 (0.03)	0.37 (0.03)	0.22 (0.02)	0.28 (0.03)	---	---	---	--	---	76.35%
0.25 (0.01)	0.59 (0.03)	0.55 (0.03)	0.37 (0.03)	0.22 (0.02)	0.28 (0.03)	0.05 (0.01)	---	---	--	---	77.16%
0.24 (0.01)	0.58 (0.03)	0.54 (0.03)	0.36 (0.03)	0.21 (0.02)	0.27 (0.03)	0.06 (0.02)	0.04 (0.01)	---	--	---	77.14%
0.24 (0.01)	0.58 (0.03)	0.54 (0.03)	0.36 (0.03)	0.22 (0.02)	0.27 (0.03)	0.06 (0.02)	0.04 (0.02)	0.09 (0.02)	0.12 (0.03)	---	78.12%
0.25 (0.01)	0.58 (0.03)	0.54 (0.03)	0.36 (0.04)	0.22 (0.02)	0.27 (0.03)	0.06 (0.02)	0.04 (0.02)	0.09 (0.02)	0.11 (0.03)	0.01 (0.01)	78.20%

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

6. Using larger values of h does not alter the results.

7. This result does not imply that the term spread is not a leading economic indicator. This implies that its leading information could be contained already in the rest of the economic indicators included in the model.

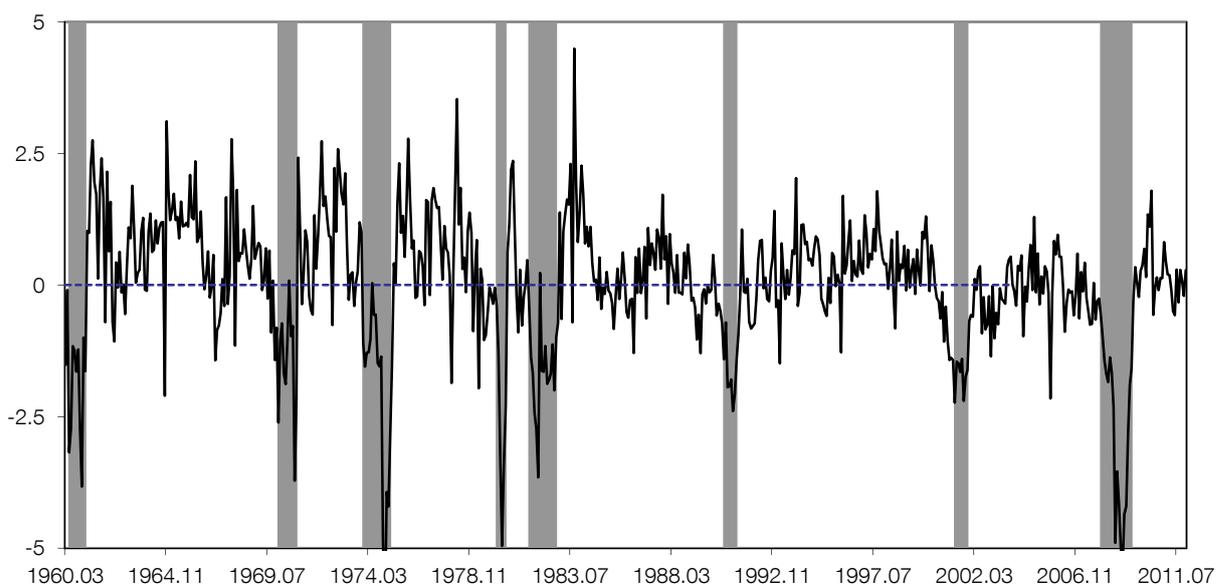
Figure 1. Log likelihood and lead time of financial indicators



Notes. SP500 and the term spread at time t have been related to the common factor at time $t+h$. In this figure, h appears in the horizontal axis and the log likelihoods reached by the dynamic factor model appear in the vertical axis.

As in the original proposal, our extension of AD is based on the notion that co-movements among the macroeconomic variables have a common element, the common factor that moves in accordance with the US business cycle dynamics. To check whether the business cycle information that can be extracted from the common factor agrees with the US business cycle, the coincident indicator along with shaded areas that refer to the NBER recessionary periods are plotted in Figure 2. The figure shows the high performance of the coincident indicator as a business cycle indicator since it is in striking accord with the professional consensus as to the history of US business cycle. During the periods that the NBER classifies as expansions, the values of the coincident indicator are usually positive. At around the beginning of the NBER-dated recessions, the common factor drastically falls and remains low until around the times the NBER dates the end of the recessions.

Figure 2. Common factor



Notes: Shaded areas correspond to recessions as documented by the NBER.

To analyze in depth the accuracy of the common factor to compute business cycle inferences, let us assume that there is a regime switch in the index itself.⁸ 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 , that is allowed to follow a first-order Markov chain. Following Hamilton (1989), a simple switching model may be specified as:

$$x_t = c_{s_t} + \sum_{j=1}^p \alpha_j x_{t-j} + \varepsilon_t \quad (16)$$

where $\varepsilon_t \sim iidN(0, \sigma)$.⁹ 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 $t-1$, \mathcal{X}_{t-1} , and of the business cycle states prior to $t-1$. Accordingly, the probabilities of staying in each state are

$$p(s_t = i / s_{t-1} = j, s_{t-2} = h, \dots, \mathcal{X}_{t-1}) = p(s_t = i / s_{t-1} = j) = p_{ij}. \quad (17)$$

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 the second regime with recessions. According to the related literature, expansions are more persistent than downturns (estimated p_{00} and p_{11} of about 0.98 and 0.91, respectively). These estimates are in line with the well-known fact that expansions are longer than contractions, on average.

Finally, Figure 3 displays the estimated smoothed probabilities of recessions along with shaded areas that refer to the periods classified as recessions by the NBER. The figure illustrates the great ability of the model to capture the US business cycle and validates the interpretation of state $s_t = 1$ as a recession and the probabilities plotted in this chart as probabilities of being in recession.

Table 3. Markov-switching estimates

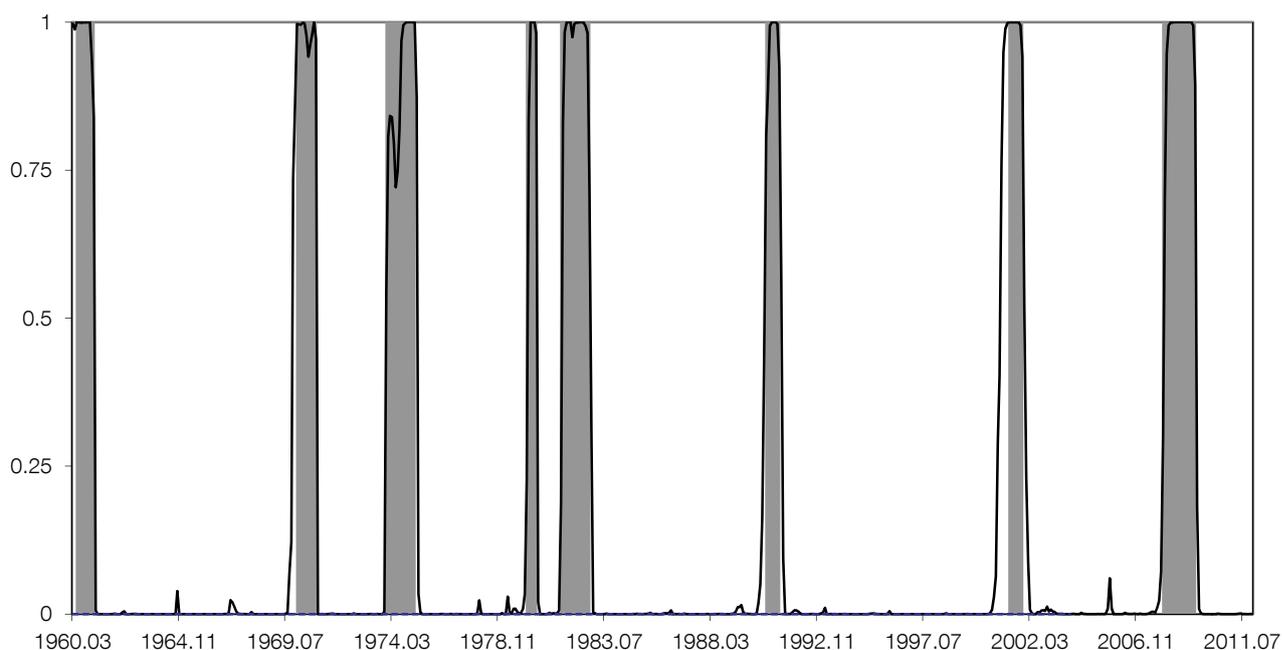
c_0	c_1	σ^2	p_{00}	p_{11}
0.38	-1.99	0.89	0.98	0.91
(0.04)	(0.11)	(0.05)	(0.01)	(0.02)

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 $p(s_t = i / s_{t-1} = j) = p_{ij}$.

8. Camacho, 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.

9. According to Camacho and Perez Quiros (2007), we included no lags in the Markov-switching specification. We checked that the resulting model is dynamically complete in the sense that the errors are white noise.

Figure 3. Smoothed probabilities from common factor



Notes: Shaded areas correspond to recessions as documented by the NBER.

3.3 Simulated real-time analysis

Among many others, Stark and Croushore (2002) suggest that the analysis of the in-sample forecasting performance of competitive models is questionable since the results can be deceptively lower when using real-time vintages. This happens because the in-sample analysis misses 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 are released at different point in time; and (iii) the real time data revisions.

However, although developing real-time data sets is conceptually simple, constructing real-time vintages is sometimes, 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). Although the method is based on successive enlargements of the latest available data set, it differs from the traditional out-of-sample analysis since, when constructing the data vintages, it takes into account the real-time data flow (and hence the publication lags). Therefore, the experiment tries to mimic as closely as possible the real-time analysis that would have been performed by a potential user of the models when forecasting, at each period of time, on the basis of different vintages of data sets. Accordingly, it is labeled as “pseudo” real-time analysis since only (i) and (ii) hold.

In our pseudo real-time analysis, the data vintages are updated twice each month, on the first day and on the fifteen day of these months. The first data vintage of our experiment refers to August 1, 1989 and the last data vintage refers to January 15, 2012. Although they are collected from the information of the latest available data set, they preserved the data release calendar that a forecaster would have faced on that day. To facilitate the creation of the data vintages, we treat the data as though they were released in blocks each fifteen days.

For example, let us illustrate the enlargement process by describing the two data vintages generated in February. According to the release lag-times described in Table 1, on February 1 the data set is enlarged with sales, whose latest figure refers to November; with industrial production, income, and housing starts, whose latest figures refer to December; and with new orders, consumer confidence, PMI, and SP500, whose latest figures refer to January. When the data vintage is updated on February 15, the data set is enlarged with employment, whose latest figure refers to January. Finally, the data vintage is updated with GDP at the beginning of February, May, August, and November, whose latest figures refer to the quarters that end in December, March, June and September, respectively.

Since the model is designed to compute short-term forecasts, in each forecasting day we compute blocks of nine-month-ahead forecasts from each data vintage. Therefore, each of these blocks incorporates forecasts that refer to the last quarter's GDP growth before its official release (backcasts), others refer to the current quarter GDP growth (nowcasts), while others refer to the next quarter's GDP growth (forecasts). Remarkably, all parameters, factors, and so forth are re-estimated for each of the pseudo real-time vintages. This procedure leads to 540 blocks of forecasts for each model considered in the analysis.

Table 4: Predictive accuracy

	Backcasts		Nowcasts		Forecasts	
	Mean Squared Errors					
AD	0.257		0.369		0.445	
	E: 0.218	R: 0.536	E: 0.243	R: 1.264	E: 0.218	R: 2.033
RW	0.404		0.500		0.504	
	E: 0.210	R: 2.531	E: 0.211	R: 2.554	E: 0.217	R: 2.569
AR	0.358		0.431		0.491	
	E: 0.208	R: 1.435	E: 0.208	R: 2.007	E: 0.209	R: 2.413
MS	0.283		0.401		0.450	
	E: 0.167	R: 1.119	E: 0.218	R: 1.697	E: 0.239	R: 1.924
TAR	0.292		0.408		0.451	
	E: 0.187	R: 1.039	E: 0.318	R: 1.004	E: 0.225	R: 2.028
Our extension	0.218		0.368		0.443	
	E: 0.182	R: 0.480	E: 0.245	R: 1.238	E: 0.220	R: 1.989
	Equal predictive accuracy tests					
Our model vs RW	0.002		0.020		0.010	
Our model vs AR	0.009		0.092		0.009	
Our model vs AD	0.019		0.917		0.742	
Our model vs MS	0.010		0.040		0.085	
Our model vs TAR	0.012		0.042		0.090	

Notes. The forecasting sample is 1989.3-2011.4. The top panel shows the Mean Squared Errors (MSE) of the dynamic factor model proposed by Aruoba and Diebold (AD, 2010), a random walk (RW), an autoregressive model (AR), a Markov-switching model (MS), a threshold autoregressive model (TAR), and our extension of AD. R and E refer to recessions and expansions periods according to NBER. The bottom panel shows the p -values of the Diebold-Mariano (DM) test of equal predictive accuracy.

The predictive accuracy of the models is examined in Table 4. The table shows the mean-squared forecast errors (MSE) of each model, which is 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 proposed by AD and our extension described in Section 2 (labeled as “our model”), four benchmark models are included in the forecast evaluation. The first model is an autoregressive model of order two (AR) which is estimated in real-time producing iterative forecasts. The second model is a random walk (RW) model whose forecasts are equal to the average latest available real-time observations. According to the AR model, the third model is an autoregressive Markov-switching model of order two (MS) as described in (16) where the drift of the US growth rate is governed by an unobserved state variable. The fourth model is a threshold autoregressive specification (TAR) of order two. Following the lines of the MS model, the TAR specification is stated as

$$y_{\tau} = c_0 I_0 + c_1 I_1 + \phi_1 y_{\tau-1} + \phi_2 y_{\tau-2} + \varepsilon_{\tau}, \quad (18)$$

Where $\varepsilon_{\tau} \sim iidN(0, \sigma^2)$, the indicator functions are

$$I_0 = \begin{cases} 1 & \text{if } y_{\tau-d} \leq g \\ 0 & \text{if } y_{\tau-d} > g \end{cases} \text{ and } I_1 = \begin{cases} 0 & \text{if } y_{\tau-d} > g \\ 1 & \text{if } y_{\tau-d} \leq g \end{cases}, \quad (19)$$

where d and g are selected for each forecasts in order to maximize the likelihood function of the model.

Note that the MSEs lead 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 other models' forecasts. To analyze whether empirical loss differences between two or more competing models are statistically significant, there are a large number of tests proposed in the literature. Among them, we focus on the pairwise test introduced by Diebold and Mariano (DM, 1995) which seems to be the most influential and most widely used test. The results are reported in the bottom panel of Table 4.

From the figures reported in Table 4, several features of the pseudo real-time forecasting analysis are noteworthy. First, when comparing the forecasts from multivariate models with those from univariate models, the MSEs show that the former clearly outperform the latter. Notably, the equal predictive accuracy tests show that the differences are statistically significant.

Second, to analyze the stability of the forecasting performance over time, Table 4 also incorporates within-recessions and within-expansion MSEs, which are computed from the cycles already identified by NBER. Although the forecasting improvements of multivariate models over univariate models appear in the two phases of the business cycle, they become especially important during the NBER recessions.

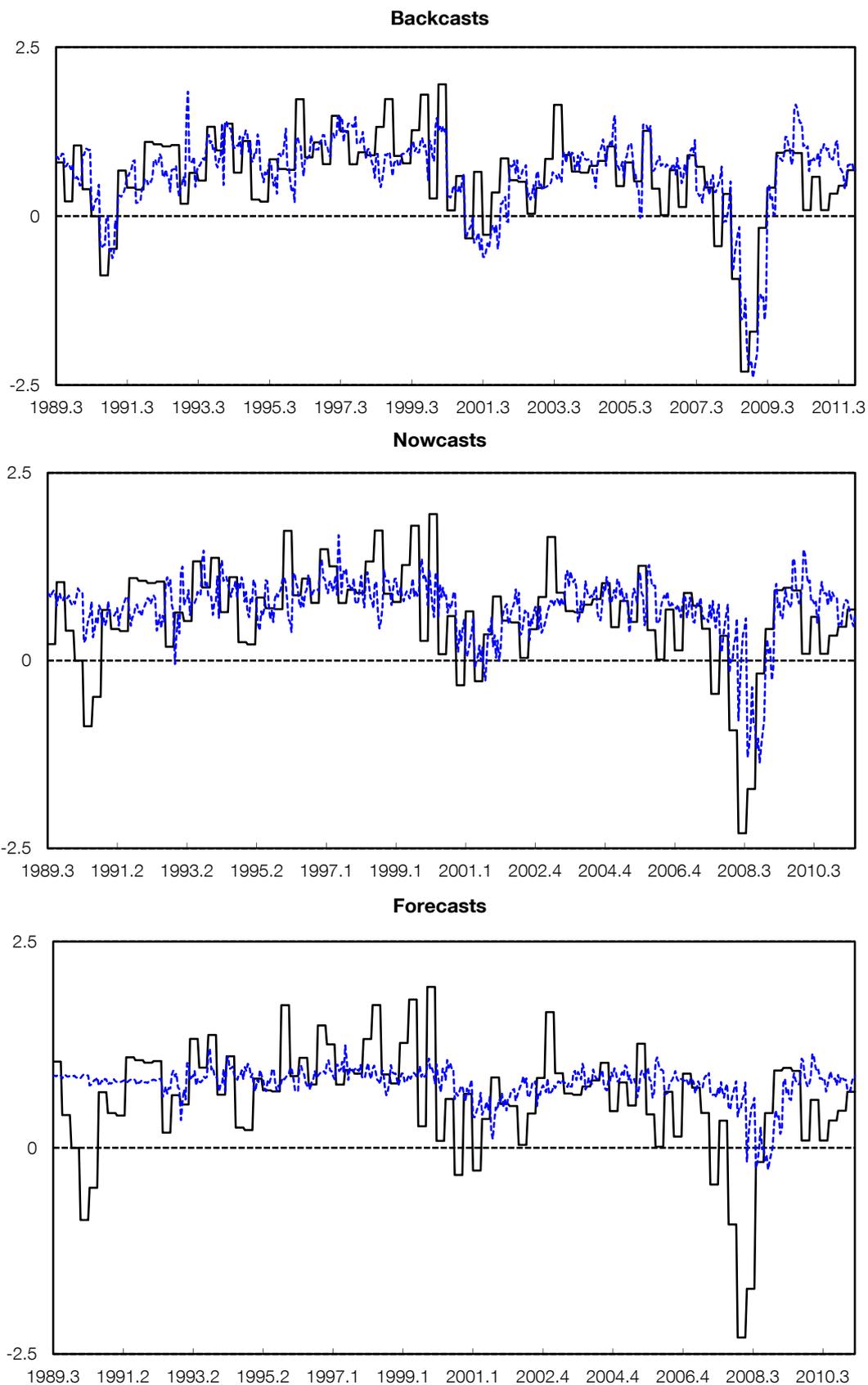
Third, the gains in using the multivariate models in forecasting GDP with respect to univariate models diminish with the forecast horizon, although they remain statistically significant. The intuition behind this result is that the factor models use the incoming information as it is available from the promptly published economic indicators. This early available information is much less valuable as the forecasting horizon increases. In fact, for large forecasting horizons the monthly indicators are not available for the reference quarter

and all the time series used in the models must be forecasted for the quarter of interest, regardless if the model is univariate or multivariate.

Finally, our extension of the Aruoba-Diebold dynamic factor model exhibits some forecast improvements over the seminal proposal. Again, the gains depend on the forecast horizon. In the backcasting exercise, the differences between the MSE results of these two factor models are noticeable (relative MSE of 0.848) and statistically significant (p -value of 0.019). This result encourages real-time forecasters that try to obtain early estimates of GDP growth during the weeks before its first release to check back at the bulk of monthly real, survey and financial data. In nowcasting and forecasting, our model still exhibits slightly lower MSEs although the gains diminish considerably (relative MSE of about 0.99) and the differences are not statistically significant. In these two cases both factor models could alternatively be used to compute the forecast.

To examine the accuracy of the forecasts visually, plots of actual and pseudo real-time predictions are shown in Figure 4. The straight lines depict simulated real-time forecasts of US GDP growth while the dashed lines refer to the corresponding final quarterly data, which are equally distributed among the respective days of the quarter for the sake of comparison. Overall, the forecasts follow sequential patterns that track the business cycle marked by the evolution of GDP releases. As expected, the real-time estimates become more accurate in the case of backcasts (top panel) since the predictions are computed immediately before the end of the quarter, which allow them to use the latest available information of the respective quarter. Accordingly, nowcasts (middle panel) and forecasts (bottom panel) track the GDP dynamics with some delays since they use poorer information sets to compute predictions although they are available sooner.

Figure 4. Real-time predictions and actual realizations



Notes. Actual realizations of GDP growth (dotted line) and real time predictions, backcasts (top), nowcasts (middle) and forecasts (bottom panel) from model M6 (see Table 2).

4 Conclusions

The dynamic factor model proposed by Aruoba and Dieblod (AD, 2010) was originally designed to produce high frequency measurement of the overall economic activity in a systematic, replicable, and statistically optimal manner from GDP, industrial production, income, sales and employment data. By means of a pseudo real-time empirical evaluation, we show that their dynamic factor model is an excellent place to compute short-term forecasts of US GDP growth in real time since it produces more accurate forecasts than several benchmarks.

In addition, we extend their model to examine the informational content of additional real activity data, survey indexes and financial indicators to compute the forecasts. According to their timely publication, we find that these indicators contain important information to compute GDP backcasts beyond the monthly real activity measures considered in AD. Since the first estimate of US quarterly GDP is released several weeks after the end of the quarter, this result is important for assessing the macroeconomic conditions in the meantime. Remarkably, nowcasts and forecasts computed from our extension are as accurate as those computed from the original AD model. Therefore, we consider that our extension is a valid tool to be used for short-term analysis.

To conclude, we consider that the work begun here could be further extended to examine the extent to which the single-index dynamic factor model is an appropriate tool to forecast other important US economic variables, such as inflation, employment, financial indicators, the GDP components, or debt.

Appendix

Without loss of generalization, we assume that our model contains only GDP, one non-financial monthly indicator and one financial monthly indicator, which are collected in the vector $Y_t = (y_t^*, Z_{it}^*, Z_{ft}^*)'$. For simplicity's sake, we also assume that $\rho_1 = \rho_2 = \rho_3 = 1$, and that the lead for the financial indicator is $h = 1$. In this case, the observation equation, $Y_t = Z\alpha_t$, is

$$\begin{pmatrix} y_t^* \\ Z_{it}^* \\ Z_{ft}^* \end{pmatrix} = \begin{pmatrix} 0 & \frac{\beta_y}{3} & \frac{2\beta_y}{3} & \beta_y & \frac{2\beta_y}{3} & \frac{\beta_y}{3} & \frac{1}{3} & \frac{2}{3} & 1 & \frac{2}{3} & \frac{1}{3} & 0 & 0 \\ 0 & \beta_i & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \beta_f & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{t+1} \\ x_t \\ \vdots \\ x_{t-4} \\ u_t^y \\ \vdots \\ u_{t-4}^y \\ u_t^i \\ u_t^f \end{pmatrix} \quad (A1)$$

It is worth noting that the model assumes contemporaneous correlation between non-financial indicators and the state of the economy, whereas for financial variables, the correlation is imposed between current values of the indicators and future values of the common factor.

The transition equation, $\alpha_t = T\alpha_{t-1} + \eta_t$, is

$$\begin{pmatrix} x_{t+1} \\ x_t \\ \vdots \\ x_{t-4} \\ u_t^y \\ \vdots \\ u_{t-4}^y \\ u_t^i \\ u_t^f \end{pmatrix} = \begin{pmatrix} \rho_1 & \cdots & 0 & 0 & 0 & \cdots & \cdots & 0 \\ 1 & & & 0 & & \cdots & & 0 \\ \vdots & \ddots & & \vdots & & \cdots & & \vdots \\ 0 & \cdots & 1 & 0 & & \cdots & & 0 \\ 0 & \cdots & & 0 & d_1^y & 0 & 0 & 0 \\ \vdots & \cdots & & \cdots & & \ddots & & \vdots \\ 0 & \cdots & & & & 1 & 0 & 0 \\ 0 & \cdots & & & & 0 & 0 & d_1^i \\ 0 & \cdots & & & & 0 & 0 & 0 & d_1^f \end{pmatrix} \begin{pmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_{t-5} \\ u_{t-1}^y \\ \vdots \\ u_{t-5}^y \\ u_{t-1}^i \\ u_{t-1}^f \end{pmatrix} + \begin{pmatrix} e_{t+1} \\ 0 \\ \vdots \\ 0 \\ \varepsilon_t^y \\ \vdots \\ \varepsilon_{t-4}^y \\ \varepsilon_t^i \\ \varepsilon_t^f \end{pmatrix}, \quad (A2)$$

where $\eta_t \sim iN(0, Q)$ and $Q = \text{diag}(\sigma_e^2, 0, \dots, 0, \sigma_y^2, 0, \dots, 0, \sigma_i^2, \sigma_f^2)$.

REFERENCES

- ARUOBA, B., DIEBOLD, F. and SCOTTI, CH. (2009), "Real-time measurement of business conditions", *Journal of Business & Economic Statistics* 27: 417-427.
- ARUOBA, B. and DIEBOLD, F. (2010), "Real-time macroeconomic monitoring: Real activity, inflation, and interactions", *American Economic Review: Pap Proc* 100: 20-24.
- BOIVIN, J. and NG, S. (2006), "Are more data always better for factor analysis?", *Journal of Econometrics* 132: 169-194.
- CAMACHO, M., PEREZ QUIRÓS, G and PONCELA, P. (2012), "Extracting nonlinear signals from several economic indicators", *CEPR Working Paper No. 8865*.
- CAMACHO, M. and DOMENECH, R. (2012), "MICA-BBVA: A factor model of economic and financial indicators for short-term GDP forecasting". *SERIES: Journal of the Spanish Economic Association* 3: 475-497.
- CAMACHO, M. and PEREZ QUIRÓS, G. (2007), "Jump-and-rest effect of U.S. business cycles", *Studies in Nonlinear Dynamics & Econometrics* 11(4): article 3.
- CAMACHO, M. and PEREZ QUIRÓS, G. (2010), "Introducing the Euro-STING: Short Term Indicator of euro area Growth", *Journal of Applied Economics* 25: 663-694.
- DIEBOLD, F. and MARIANO, R. (1995), "Comparing predictive accuracy", *Journal of Business & Economic Statistics* 13: 253-263.
- GIANNONE, D., REICHLIN, L., and SMALL, D. (2008), "Nowcasting: The real-time informational content of macroeconomic data", *Journal of Monetary Economics* 55: 665-676.
- HAMILTON, J. (1989), "A new approach to the economic analysis of nonstationary time series and the business cycles", *Econometrica* 57: 357-384.
- MARIANO, R. and MURASAWA, Y. (2003), "A new coincident index of business cycles based on monthly and quarterly series", *Journal of Applied Economics* 18: 427-443.
- STARK, T. and CROUSHORE, D. (2002), "Forecasting with Real-Time Data Set for Macroeconomists", *Journal of Macroeconomics* 24: 507-531.
- STOCK J. and WATSON M. (1991), "A probability model of the coincident economic indicators. In: Lahiri K, Moore G (eds.) *Leading economic indicators, new approaches and forecasting records*", Cambridge University Press, Cambridge.
- WHEELLOCK D. and WOHAR M. (2009), "Can the term spread predict output growth and recessions? A survey of the literature", *Federal Reserve Bank of St. Louis Rev* 91: 419-440.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

- 1301 JAMES COSTAIN and ANTON NAKOV: Logit price dynamics.
- 1302 MIGUEL GARCÍA-POSADA: Insolvency institutions and efficiency: the Spanish case.
- 1303 MIGUEL GARCÍA-POSADA and JUAN S. MORA-SANGUINETTI: Firm size and judicial efficacy: evidence for the new civil procedures in Spain.
- 1304 MAXIMO CAMACHO and GABRIEL PEREZ-QUIROS: Commodity prices and the business cycle in Latin America: living and dying by commodities?
- 1305 CARLOS PÉREZ MONTES: Estimation of regulatory credit risk models.
- 1306 FERNANDO LÓPEZ VICENTE: The effect of foreclosure regulation: evidence for the US mortgage market at state level.
- 1307 ENRIQUE MORAL-BENITO and LUIS SERVEN: Testing weak exogeneity in cointegrated panels.
- 1308 EMMA BERENQUER, RICARDO GIMENO and JUAN M. NAVE: Term structure estimation, liquidity-induced heteroskedasticity and the price of liquidity risk.
- 1309 PABLO HERNÁNDEZ DE COS and ENRIQUE MORAL-BENITO: Fiscal multipliers in turbulent times: the case of Spain.
- 1310 SAMUEL HURTADO: DSGE models and the Lucas critique.
- 1311 HENRIQUE S. BASSO and JAMES COSTAIN: Fiscal delegation in a monetary union with decentralized public spending.
- 1312 MAITE BLÁZQUEZ CUESTA and SANTIAGO BUDRÍA: Does income deprivation affect people's mental well-being?
- 1313 ENRIQUE ALBEROLA, ÁNGEL ESTRADA and DANIEL SANTABÁRBARA: Growth beyond imbalances. Sustainable growth rates and output gap reassessment.
- 1314 CARMEN BROTO and GABRIEL PEREZ-QUIROS: Disentangling contagion among sovereign CDS spreads during the European debt crisis.
- 1315 MIGUEL GARCÍA-POSADA and JUAN S. MORA-SANGUINETTI: Are there alternatives to bankruptcy? A study of small business distress in Spain.
- 1316 ROBERTO RAMOS and ENRIQUE MORAL-BENITO: Agglomeration matters for trade.
- 1317 LAURA HOSPIDO and GEMA ZAMARRO: Retirement patterns of couples in Europe.
- 1318 MAXIMO CAMACHO, GABRIEL PEREZ-QUIROS and PILAR PONCELA: Short-term forecasting for empirical economists. A survey of the recently proposed algorithms.
- 1319 CARLOS PÉREZ MONTES: The impact of interbank and public debt markets on the competition for bank deposits.
- 1320 OLYMPIA BOVER, JOSE MARIA CASADO, SONIA COSTA, PHILIP DU CAJU, YVONNE MCCARTHY, EVA SIERMINSKA, PANAGIOTA TZAMOURANI, ERNESTO VILLANUEVA and TIBOR ZAVADIL: The distribution of debt across euro area countries: the role of Individual characteristics, institutions and credit conditions.
- 1321 BRINDUSA ANGHIEL, SARA DE LA RICA and AITOR LACUESTA: Employment polarisation in Spain over the course of the 1997-2012 cycle.
- 1322 RODOLFO G. CAMPOS and ILIANA REGGIO: Measurement error in imputation procedures.
- 1323 PABLO BURRIEL and MARÍA ISABEL GARCÍA-BELMONTE: Meeting our D€STINY. A Disaggregated €uro area Short Term Indicator model to forecast GDP (Y) growth.
- 1401 TERESA SASTRE and FRANCESCA VIANI: Countries' safety and competitiveness, and the estimation of current account misalignments.
- 1402 FERNANDO BRONER, ALBERTO MARTIN, AITOR ERCE and JAUME VENTURA: Sovereign debt markets in turbulent times: creditor discrimination and crowding-out effects.
- 1403 JAVIER J. PÉREZ and ROCÍO PRIETO: The structure of sub-national public debt: liquidity vs credit risks.
- 1404 BING XU, ADRIAN VAN RIXTEL and MICHIEL VAN LEUVENSTEIJN: Measuring bank competition in China: a comparison of new versus conventional approaches applied to loan markets.
- 1405 MIGUEL GARCÍA-POSADA and JUAN S. MORA-SANGUINETTI: Entrepreneurship and enforcement institutions: disaggregated evidence for Spain.
- 1406 MARIYA HAKE, FERNANDO LÓPEZ-VICENTE and LUIS MOLINA: Do the drivers of loan dollarisation differ between CESEE and Latin America? A meta-analysis.
- 1407 JOSÉ MANUEL MONTERO and ALBERTO URTASUN: Price-cost mark-ups in the Spanish economy: a microeconomic perspective.
- 1408 FRANCISCO DE CASTRO, FRANCISCO MARTÍ, ANTONIO MONTESINOS, JAVIER J. PÉREZ and A. JESÚS SÁNCHEZ-FUENTES: Fiscal policies in Spain: main stylised facts revisited.

- 1409 MARÍA J. NIETO: Third-country relations in the Directive establishing a framework for the recovery and resolution of credit institutions.
- 1410 ÓSCAR ARCE and SERGIO MAYORDOMO: Short-sale constraints and financial stability: evidence from the Spanish market.
- 1411 RODOLFO G. CAMPOS and ILIANA REGGIO: Consumption in the shadow of unemployment.
- 1412 PAUL EHLING and DAVID HAUSHALTER: When does cash matter? Evidence for private firms.
- 1413 PAUL EHLING and CHRISTIAN HEYERDAHL-LARSEN: Correlations.
- 1414 IRINA BALTEANU and AITOR ERCE: Banking crises and sovereign defaults in emerging markets: exploring the links.
- 1415 ÁNGEL ESTRADA, DANIEL GARROTE, EVA VALDEOLIVAS and JAVIER VALLÉS: Household debt and uncertainty: private consumption after the Great Recession.
- 1416 DIEGO J. PEDREGAL, JAVIER J. PÉREZ and A. JESÚS SÁNCHEZ-FUENTES: A toolkit to strengthen government budget surveillance.
- 1417 J. IGNACIO CONDE-RUIZ, and CLARA I. GONZÁLEZ: From Bismarck to Beveridge: the other pension reform in Spain.
- 1418 PABLO HERNÁNDEZ DE COS, GERRIT B. KOESTER, ENRIQUE MORAL-BENITO and CHRISTIANE NICKEL: Signalling fiscal stress in the euro area: a country-specific early warning system.
- 1419 MIGUEL ALMUNIA and DAVID LÓPEZ-RODRÍGUEZ: Heterogeneous responses to effective tax enforcement: evidence from Spanish firms.
- 1420 ALFONSO R. SÁNCHEZ: The automatic adjustment of pension expenditures in Spain: an evaluation of the 2013 pension reform.
- 1421 JAVIER ANDRÉS, ÓSCAR ARCE and CARLOS THOMAS: Structural reforms in a debt overhang.
- 1422 LAURA HOSPIDO and ENRIQUE MORAL-BENITO: The public sector wage premium in Spain: evidence from longitudinal administrative data.
- 1423 MARÍA DOLORES GADEA-RIVAS, ANA GÓMEZ-LOSCOS and GABRIEL PÉREZ-QUIRÓS: The Two Greatest. Great Recession vs. Great Moderation.
- 1424 ENRIQUE MORAL-BENITO and OLIVER ROEHN: The impact of financial (de)regulation on current account balances.
- 1425 MÁXIMO CAMACHO and JAIME MARTÍNEZ-MARTÍN: Real-time forecasting US GDP from small-scale factor models.