NOWCASTING SPANISH GDP GROWTH IN REAL TIME: "ONE AND A HALF MONTHS EARLIER"

2010

David de Antonio Liedo and
Elena Fernández Muñoz

Documentos de Trabajo
N.º 1037

BANCO DE ESPAÑA
Eurosistema

NOWCASTING SPANISH GDP GROWTH IN REAL TIME: “ONE AND A HALF MONTHS EARLIER”

NOWCASTING SPANISH GDP GROWTH IN REAL TIME: “ONE AND A HALF MONTHS EARLIER”

David de Antonio Liedo (*) and Elena Fernández Muñoz

BANCO DE ESPAÑA

(*) Corresponding author (david.de.antonio.liedo@gmail.com). This paper has benefited from stimulating discussions with several members of the “Economic Analysis and Forecasting Department” at the Banco de España. We would also like to thank, without implicating, Samuel Hurtado, Gabriel Perez-Quiros and Alberto Urtasun for comments and suggestions at the earliest stage of the work. *DISCLAIMER: The views expressed in this paper are the author’s, not those of Banco de España.*

Documentos de Trabajo. N.º 1037

2010

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, 2010

ISSN: 0213-2710 (print)
ISSN: 1579-8666 (on line)
Depósito legal: M. 52842-2010
Unidad de Publicaciones, Banco de España

Abstract

The sharp decline in economic activity registered in Spain over 2008 and 2009 has no precedents in recent history. After ten prosperous years with an average GDP growth of 3.7%, the current recession places *non-judgemental* forecasting models under *stress.* This paper evaluates the Spanish GDP *nowcasting* performance of combinations of small and medium-sized linear dynamic regressions with priors originating in the Bayesian VAR literature. Our forecasting procedure can be considered a timely and simple approximation to the mix of accounting tools, models and judgement used by the statistical agencies to construct aggregate GDP figures. The real time forecast evaluation conducted over the most severe phase of the recession shows that our method yields reliable real GDP growth predictions almost *one and a half months* before the official figures are published.

Keywords: Minnesota priors, mixed estimation, forecasting.

JEL classification: C32, C53, E37.

Resumen

El fuerte descenso de la actividad económica registrado en España durante 2009 y 2010 no tiene precedentes en la historia más reciente. Tras diez años de prosperidad con un crecimiento medio del 3,7%, el escenario macroeconómico actual somete a estrés los modelos de predicción automaticos. En este artículo se evalua la capacidad de varias combinaciones de modelos multivariantes autoregresivos con retardos distribuidos (ADL) para obtener “nowcasts” o estimaciones del PIB anteriores a la publicación oficial. Dichos modelos requiren la estimación de un elevado número de parámetros cuando desea construirse una predicci´on condicional a un amplio conjunto de variables. Para hacer frente a la llamada “maldición de la dimensionalidad”, utilizamos información a priori proveniente de la literatura sobre Vectores Autorregresivos Bayesianos (BVAR). Nuestro procedimiento puede interpretarse como un método simple y oportuno para aproximar la mezcla de herramientas contables, modelos y juicio que se utiliza en cualquier agencia estadística durante el proceso de construccion de las cifras del PIB agregado. La evaluación en tiempo real durante la fase más severa de la actual recesión muestra que nuestro método permite obtener predicciones fiables del PIB real español casi un mes y medio antes de que las cifras oficiales se hagan públicas.

Palabras clave: Vectores autorregresivos bayesianos, estimación mixta, predicción.

Códigos JEL: C32, C53, E37.

1 Introduction

After ten years of stable growth slightly below 4% per year in Spain, the current recession provides an excellent opportunity to "*stress-test*" our *non-judgemental* forecasting models in real time. The use of real-time data for model validation simulates the actual environment of professional forecasters and, as suggested by Stark and Croushore (2002), avoids misleading conclusions that may be obtained when the models are estimated and used on the basis of lastest available data.

As defined by Giannone et al. (2008) or Banbura et al. (2010b), *nowcasting* refers to the prediction of the most recent past, the present, and the nearest future[1]. In this paper, we conduct a *post-mortem* exercise to evaluate the accuracy with which Spanish GDP growth can be predicted using information *sub*sets available to the forecasters one and a half months before the official GDP figure is released by the statistical agency[2]. Our nowcasts take the form of a linear combination of largely unrestricted regression equations that aim to approximate the performance of the mixture of models and judgment used by the statistical agency to construct the GDP figures.

The existing tools available for nowcasting Spanish GDP growth in real time take into account the presence of strong co-movements in macroeconomic data by incorporating restrictions inspired by the literature on dynamic factor models, e.g. Camacho and Domenech (2010), Camacho and Pérez-Quirós (2010b), Cuevas and Quillis (2010). Factor models are relatively restrictive representa-

[1] Although these papers use the general term *"nowcasting"* regardless of whether the aim is to predict the last, current or next quarter, some papers in the literature, e.g. Angelini et al. (2010), distinguish between backcasting, nowcasting and forecasting. This would imply that predictions for the first quarter based on the information set available on December 31st would be called forecasts, while predictions constructed one day later, on January 1st, would be nowcasts.

[2] The web site of the Spanish National Statistical Agency ("I.N,.E." by its Spanish acronym) can be found at www.ine.es.

tions which allow GDP growth to be expressed as the sum of two orthogonal components: one driven by pervasive factors that spread throughout the economy, and a measurement error component that is idiosyncratic. Such restrictions have also been successful in nowcasting US and euro area data, as shown by Giannone et al.(2008) and Angelini et al. (2010), respectively.

Our projections conditional on the available predictor variables are based on dynamic regression models (autoregressive distributed lags). As opposed to the class of forecasting tools mentioned above, we do not *impose* the presence of co-movements by shrinking the available monthly information into one or a few quarterly factors. The potential multicollinearity problems arising from the large amount of synchronization among the predictor variables is offset by the use of priors or "inexact" restrictions originated in the VAR literature. Interestingly, De Mol et al. (2008) show that forecasts based on large Bayesian (static) regressions can be highly correlated with those resulting from *static* principal components. Thus, our *dynamic* regressions have the potential to capture the business cycle co-movements without having to impose a *dynamic* factor analytical structure. The large and medium-sized Bayesian VARs developed by Banbura et al. (2010a) to forecast monthly US macro variables illustrate this idea and help to motivate the use of dynamic regressions also in the field of nowcasting.

To our knowledge, our paper presents the first real-time "nowcasting" exercise with medium-sized Bayesian dynamic regression models. In general, the larger the number of indicators included in a regression, the smaller the risk of model misspecification. The models we consider allow us to obtain GDP projections conditional on the first p lags and the current and past values of a set of N indicator variables. This requires the estimation of a very large number of parameters, which could lead to in-sample overfitting and large out-of-sample forecast errors. However, the use of Minnesota-type priors on VAR coefficients as a method of tackling the curse of dimensionality has been standard practice since Litterman (1980, 1986), and it has been shown to be a valid strategy even when the number of variables is large (see Banbura et al. 2010a).

Our real-time forecasting exercise features two additional innovations. First, we illustrate the potential advantages of defining the prior for a given forecasting equation with an Empirical Bayes method (Robbins, 1954) that uses the data to determine the prior. We grant a higher hierarchical level to the parameters defining the prior's shrinkage than to the regression coefficients, and identify its most likely values from a pre-sample or *training sample*. A similar approach is followed by Giannone et al. (2010), who propose to set the prior hyperparameters to values that maximize the marginal likelihood of the data in the context of VAR models.

A second key element of our approach is that we take into account model uncertainty. Each one of the models considered allows us to construct a projection conditional on a particular subset of indicators. An information set based on N predictor variables yields a total of $2^N - 1$ different nowcasts for real GDP. Thus, the set of models can be represented by $M = \{M_1, M_2, \ldots, M_{2^N-1}\}$. Although it is common among bayesian econometricians to assume that only one of the $2^N - 1$ forecasting equations corresponds to the actual data generating process, it is typical to find posterior model probabilities that do not favour any particular model. This leads us to explore simple forecast combination strategies that attribute more weight to the models with the smallest forecast errors throughout the training sample or, alternatively, equal weights for all models.

The idea of combining models is motivated here by the real-time nature of the nowcasting problem. In real time, it is very hard to justify the use of a particular model, whereas ex-post it is possible to find models that *would have* yielded accurate forecasts. Finding a model that performs well ex-post does not provide any evidence on the real-time predictability of GDP growth. Moreover, the advantages of forecast combinations have been widely explored in forecasting applications since Bates and Granger (1969). Several papers have applied this idea to macroeconomic data and found that the best ex-ante individual forecasting models are outperformed by simple combination strategies[3]. Although the reasons are still not well understood, the literature has considered alternative explanations, often related to structural changes or non-linearities in the data generating process: see Timmermann (2006) and references therein for a review. We will argue that the success of our forecast combination of medium-sized forecasting models is based on the same principles as the success of the factor models: considerable comovements over the business cycle, and the presence of measurement errors. Although the use of only one model for real-time forecasting may be subject to criticism, it allows to determine precisely how each one of the indicators contributes to forecast GDP (see Banbura and Runstler, 2010). As shown by Banbura and Modugno (2010) and Banbura et al. (2010b), a single model can help analyze the informative content provided by intraquarterly publication of "news". This type of analysis is not required in our case, since we obtain only one nowcast per quarter, i.e. one and a half months earlier the official publication. However, our method could be extended to allow for a coherent news analysis in the presence of model uncertainty.

This paper is structured as follows. Section 2 defines the information structure available one and a half months before the official GDP figure is published and relates our forecasting method to some state-of-the art forecasting models currently available for the Spanish economy. Section 3 describes the key features of our real-time forecasting exercise, including the prior elicitation and model combination strategies. Section 4 provides the empirical results, as well as an evaluation of alternative *ex-ante* forecasting strategies and a comparison with a survey of profesional forecasters. The last section concludes.

[3]See for example Garcia-Ferrer et al. (1987), Stock and Watson (1999), Stock and Watson (2004), Andersson and Karlsson (2008), Eklund and Karlsson (2007) or Clark and McCracken (2010)

2 “Nowcasting” Spanish GDP Growth with Real-Time Data

The nowcasting problem is illustrated in Table 1. Consider, for example, the information available at the beginning of July 2010. Approximately one and a half months before the official GDP release is published by the statistical agency, monthly employment figures and various surveys corresponding to April, May and June are available. Other important variables such as sales and industrial production are available only for April and May. Finally, real exports and imports are available only for April, the first month of the previous quarter. The complete list of variables used can be found in Table 2.

Table 1: The Nowcasting Problem

Information available on 5th July 2010		PIB real	Real Exports	Real Imports	Industrial Production	Sales	Employment	Condifence in Trade	PMI Index
. . .		. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
jan-10		1,92922E+11	12985	19007	80,7	59521	17775098	-20,5	48,83
feb-10			13227	18220	81,4	59750	17741445	-11,2	47,12
mar-10			14637	19260	82,9	60061	17714299	-11,8	51,27
apr-10	NOWCASTING: previous quarter	?	13046	18828	76,6	56637	17692683	-9,5	50,93
may-10					83,0	60396	17669463	-15,1	52,26
jun-10							17645593	-14,4	51,75
jul-10	NOWCASTING: current quarter	?							
aug-10									
sep-10									
oct-10	NOWCASTING: next quarter	?							
nov-10									
dec-10									

This information can be exploited to *estimate* real GDP growth almost one and a half months before the statistical agency (I.N.E) publishes the official release.

Table 2: Data availability one week after the end of each quarter

		Start	Months Available	Source	Download
1	Consumer Confidence Indicator	1990m1	3 months	European Comission	http://ec.europa.eu/economy_finance/db_indicators/index_en.htm
2	Retail Trade Confidence Indicator	1990m1	3 months		
3	Industrial Confidence Indicator	1990m1	3 months		
4	Economic Sentiment Indicator	1990m1	3 months		
5	PMI Services	1998m2	3 months	MARKIT	http://www.markiteconomics.com/MarkitFiles/Pages/PressCenter.aspx
6	PMI Industry	1998m2	3 months		
7	Construction Employment	2001m1	3 months	Ministery of Labour	http://www.mtas.es/es/estadisticas/mercado_trabajo/index.htm
8	Total Employment	2001m1	3 months		
9	Car Registrations	1990m1	3 months	DGT	http://www.dgt.es/portal/es/seguridad_vial/estadistica/
10	IBEX'35 (Stock Exchange Index)	1990m1	3 months	Bank of Spain	http://www.bde.es/webbde/es/estadis/infoest/sindi.html
11	Industrial Production Index (non-energy)	1990m1	2 months	INE	http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft05/p050&file=inebase&L=0
12	Hotel Stays by foreigners	1990m1	2 months	INE (Encuesta de ocupación hotelera)	http://www.ine.es/jaxi/menu.do?L=0&type=pcaxis&path=%2Ft11%2Fe162eoh&file=inebase
13	Sales (non-financial)	1996m1	2 months	Tax Office	http://www.aeat.es/wps/portal
14	Sales (big firms)	1996m1	2 months		
15	Air Transportation. (Metric Tones)	1990m1	1 month	Ministery of Public Works (DG Civil Aviation)	http://www.fomento.es/BE/?nivel=2&orden=03000000
16	Building Permits	1991m11	1 month	Ministery of Public Works	http://www.fomento.es/MFOM/LANG_CASTELLANO/INFORMACION_MFOM/INFORMACION_ESTADISTICA/Construccion/
17	Real Exports	1990m1	1 month	Customs	http://serviciosweb.meh.es/APPS/DGPE/BDSICE/Busquedas/Busquedas.aspx
18	Real Imports	1990m1	1 month		
19	Imported Oil Price in Euros	1990m1	1 month	Ministery of Economics	http://serviciosweb.meh.es/APPS/DGPE/BDSICE/Busquedas/Busquedas.aspx
20	**Real GDP**	**1995q1**	**-**	**INE**	**http://www.ine.es/jaxi/menu.do?type=pcaxis&path=%2Ft35%2Fp009&file=inebase&L=0**

2.1 Our Method

The data set that is relevant for calculating a given nowcast for GDP features a "jagged edge" or missing observations at the end of the sample for some variables. Moreover, since GDP is a quarterly variable that is given by the flow of economic transactions over three months, it is possible to estimate a latent monthly GDP with interpolation methods (e.g. Chow and Lin, 1971). This type of approach is used in the context of dynamic factor models by Banbura and Modugno (2010) and by Camacho and Pérez-Quirós (2010a), who define quarterly GDP as a linear combination of unobserved factors. Giannone et al. (2010) follow a similar approach in the context of VAR models that aim to exploit the timeliness of several surveys produced by the European Commission.

The strategy followed in this paper circumvents the problem of extracting a monthly GDP signal. We simply transform all available monthly indicators into quarterly variables by using a simple aggregation rule to *bridge* quarterly GDP with monthly information. This method is simple and has the potential advantage of reducing the noise in the monthly information. A similar approach is followed by Giannone et al. (2008), who propose to *bridge* GDP growth with quarterly factors extracted from monthly indicators. As recently suggested by Armesto et al. (2010), this is a valid strategy to mix frequencies. Evaluation of alternative approaches based on the interpolation or Kalman filtering methods is left for future research.

For a given subset of N variables, the nowcast for our variable of interest Y_t is given by a simple linear projection on all available predictors and its lags:

$$\begin{aligned} P\left(Y_t|\Omega\right) = \hat{a}_1 \quad &+ \quad \hat{b}_{11} Y_{t-1} \; + \; \sum_{i=1}^{N} \hat{b}_{1,1+i} \; X_{i,t} \\ &+ \quad \hat{c}_{11} Y_{t-2} \; + \; \sum_{i=1}^{N} \hat{c}_{1,1+i} \; X_{i,t-1} \\ &+ \quad \ldots \\ &+ \quad \hat{d}_{11} Y_{t-p} \; + \; \sum_{i=1}^{N} \hat{d}_{1,1+i} \; X_{i,t-p+1} \quad + \; e_t \end{aligned} \tag{1}$$

where Ω represents the available information set and $X_{i,t}$ is the value of a given indicator i averaged over the last available three months. The symbol $\hat{}$ above the parameters indicates that they have been identified with the mixed estimation approach of Theil and Goldberger (1961). That is, sample information is mixed with dummy observation priors that reflect the presence of unit roots in the data.

Therefore, our forecasting equation has the form of a *multivariate filter* that aims to identify from the N indicators the signal revealing the most likely realization of real GDP, conditional on any given information *sub*sets.

Dummy Observation Priors

Our largely parameterized autoregressive distributed lag models will be estimated with priors originating in the BVAR literature. In paticular, we combine the Minnesota-type prior (see Litterman 1984) with priors that take into account the degree of persistence and cointegration in the variables. Those priors are parameterized here through τ, λ, μ, and d, following the notation of Lubik and Schorfheide (2005). The hyperparameter τ is the *overall tightness* of the prior. This hyperparameter helps to define the inertial behavior of the log-level of real GDP by shrinking towards one the coefficient associated with its first lag. The so-called *co-persistency prior* is introduced separately and controlled through the hyperparameter λ. This prior was originally defined by Sims (1993) as a"dummy initial observation", giving plausibility to the presence of a single stochastic trend behind the non-stationarity of the data. Another prior that serves to shrink the parameter estimates of our regression is the *own-persistence prior*, which is also known a the "sum of coefficients prior" (see Sims and Zha, 1998). The tightness of this prior is given by the hyperparameter μ. This prior, which has been proven to be useful in the framework of Large Bayesian VARs (see Banbura et al., 2010a), represents the belief that there is a unit root in each series and and weak co-movements *at very low frequencies* (no co-integration). Finally, the prior that shrinks towards zero the coefficients associated with lagged variables is governed by the hyperparameter d. The details on the implementation of the priors through dummy observations are explained in Appendix A.2.

2.2 Comparison of Alternative "Nowcasting" Methods

A comparison of several nowcasting methods currently in use for nowcasting Spanish GDP growth will clarify the added value of our approach (see Table 3).

MICA (Camacho and Domenech, 2010), Spain-Sting (Camacho and Perez-Quirós, 2010b), and FASE (Cuevas y Quillis, 2010) take into account the presence of strong co-movements in macroeconomic data by summarizing all monthly indicators in terms of one pervasive factor. This factor is behind all the co-movements we have observed during the current recession.

Table 3: Comparison of methods to "nowcast" Spanish GDP growth in real-time

			Out-of-Sample EVALUATION		
alternative methods	Number of variables	Requires data stationarity / seasonal adjustment	Period	Predictions based on *real-time/revised vintages*	Predictions based on the *real-time information* structure
Our Method	20	no/yes	2006Q3 - 2010Q2	real-time	yes
Spain-Sting	10	yes/no	2008Q1 - 2008Q4	real-time	yes
MICA-BBVA	12	yes/yes	1999Q1 - 2009Q1	revised	yes
FASE	32	yes/yes	2006Q1 - 2009Q4	revised	yes*

* The real-time use of FASE is illustrated for 2009Q4, which is the last quarter available in their sample.

In contrast, our method combines projections based on largely unrestricted dynamic regression models represented by equation 1. If GDP growth is largely driven by a single shock or factor, in line with the papers mentioned above, our benchmark equation should be able to identify it as a linear combination of current and past values of the indicators. As shown by De Mol et al. (2008), forecasts based on large Bayesian (static) regressions can be highly correlated with those resulting from *static* principal components. Thus, it makes sense to consider that our large *dynamic* regressions have the potential to capture the business cycle co-movements without any need to impose a *dynamic* factor analytical structure. The large and medium-sized Bayesian VARs developed by Banbura et al. (2010a) to forecast monthly US macro variables illustrate this idea and motivate the use of dynamic regressions also in the field of “nowcasting”.

Evaluation: Out-of-Sample Forecast Accuracy in Real-Time

Although all the models represented in Table 3 focus on the recession episode that started in 2008, not all of them are evaluated over the same sample. The longest evaluation period corresponds to the MICA-BBVA model. However, the data used to estimate the model and to construct the projections is not based on the series available in real time. As suggested by Stark and Croushore (2002), the forecasting evaluation can be misleading when latest available data is used instead of real-time data. The validation proposed by the authors of the FASE model is purely based on the real-time information structure for 2009Q4. However, the out-of-sample experiment proposed in the paper for the 2006-2009 period considers an estimate of the unobserved factor conditional on full sample information. This is not a minor detail, since the conditional expectation of the time series of unobserved factors is likely to undergo significant revisions in real time[4].

The only methods evaluated purely in real time are actually Spain-Sting and

[4] The factors are specified as a time series of unobserved variables whose expectation conditional on the information set available is obtained with the Kalman Filter.

our model combination approach. Thus, our nowcasts can only be directly compared with the Spain-Sting predictions obtained a few days after the end of each quarter[5]. Spain-Sting estimated a negative growth rate for the first quarter of 2008, while our model combination-based estimates were above 0.5%. The first official figure for that quarter was 0.27%, which is approximately the average between our estimate and the one given by Spain-Sting. However, this figure was revised upward in the official release that took place in August 2010. In addition to that, the statistical agency revised the second and third quarters of 2008 downward. Although 2008Q2 is better anticipated by Spain-Sting, this model underestimates the magnitude of the large drop in economic activity that took place in the subsequent quarter, which is actually anticipated by our model combination approach. Finally, both approaches are equally accurate at forecasting the last quarter of 2008.

[5]The forecasting performance of both approaches over 2008 can be compared on the basis of Figure 5 in their paper (Camacho and Perez-Quiros, 2010b) and Figure 8 in our paper.

3 Design of the Forecasting Exercise

This paper illustrates the real-time nature of the nowcasting problem. The choice of predictor variables and modeling strategies in real time is not straightforward to reproduce. With the benefit of hindsight, we know that monthly employment figures would have been very useful for nowcasting the gradual deceleration of 2007 and the strong GDP decline that took place in 2008Q2 and 2008Q3 (see Figure 10). However, these two quarters were subject to a large amount of uncertainty[6], and *real-time forecasters* were closely monitoring many other variables in order to understand the expected magnitude of the decline in growth.

Also with the benefit of hindsight, one could select the model that would have rendered the most accurate projections among the millions of models available. Nevertheless, the practice of real-time forecasting requires the use of an *ex-ante* strategy to determine which models to use and how to combine them. In this paper, we reproduce ex-ante strategies for nowcasting in real-time in a "simplified" context where thousands of models are available.

3.1 Real-Time Data

Seasonally adjusted GDP is obtained directly from the OECD real-time database[7]. The database contains the National statistical agency's releases since 1995 (Base 2000). A real-time database with GDP figures earlier than 1995 does not exist. Extending the database with older vintages along the lines of Croushore and Stark (2001) could be very useful to evaluate the performance of alternative forecasting methods over previous recessions.

The real-time nature of the forecasting practice determines the design of our evaluation exercise. The indicators described in Figure 2 will be seasonally ad-

[6]The statistical agency itself has announced in August 2010 a significant downward revision of the 2008Q3 GDP figure initially published more than one year ago.

[7]See http://stats.oecd.org/mei/default.asp?rev=1

justed *in real-time* using TRAMO-SEATS[8], and introduced as predictor variables in equation (1) defined in the previous section. Notice that some of our time series are quite short. Employment figures, for example, start very recently, in 2001. As opposed to the older series, which describe the number of employed individuals registered at the end of the month, the current series present the average employment registrations of each month.

Except for the confidence indicators, which enter the models without any transformation, all variables are expressed in log-levels[9].

3.2 Prior Elicitation

In this paper two alternative ways of defining the precision parameters associated with the priors are evaluated.

An Empirical Bayes Approach (EB)

Rather than using subjective beliefs, Empirical Bayes (EB) methods (Robbins, 1954) use *sample* information to elicit the priors. We explore here a method in this vein in order to choose the values of the hyperparameters defined in Subsection 2.1 (see Appendix A.2. for further details). Thus, we use a training sample to evaluate out-of-sample forecast accuracy and select the value of $h^* = [\tau^*, \lambda^*, \mu^*, d]$ that yields the most precise forecast in terms of root mean square error (RMSE). The average values of the so-called hyperparameters are given in Figure 1 as a function of the model size.

[8] Software developed at the Bank of Spain. See references and downloading options at: http://www.bde.es/webbde/es/secciones/servicio/software/econom.html

[9] An alternative to the use of TRAMO-SEATS could be to take the models directly to the raw data with Seasonal BVARs like those developed by Raynauld and Simonato (1993). A Matlab Library with a simple implementation of Seasonal BVAR models has been written by E. Quillis in www.mathworks.com. Evaluating the empirical success of this alternative option is left for future research.

A similar approach is followed by Giannone et al. (2010), in the context of VAR models. These authors propose to choose the priors that maximize the multivariate marginal likelihood of the data, which is equal to the integral of the likelihood over the prior probability measure. In the context of *large* BVARs, Banbura et al. (2010a) select the priors that yield a desired in-sample fit, in a clear effort to avoid in-sample overfitting. Both strategies admit a higher level of hierarchy for the parameters defining the prior shrinkage than for the regression coefficients.

Our strategy can be interpreted in a straightforward way. If we think of each value of h as one model, the optimal value h^* can be considered as the best forecasting model over the training sample. This implies that our out-of-sample projections would have been very precise over the training sample if the value of h^* had been "revealed" to us ex-ante.

Diffuse Priors (DP)

An important drawback of the Empirical Bayes approach outlined above is that the resulting prior for larger models can be too tight if the training sample is dominated by a period of stable growth[10]. In this case, our prior optimization results in models in which GDP growth reacts smoothly to fluctuations in indicator variables. This efficient behaviour helps over such training sample, but it comes at the cost of overpredicting GDP growth in periods of time when all indicators suddenly drop.

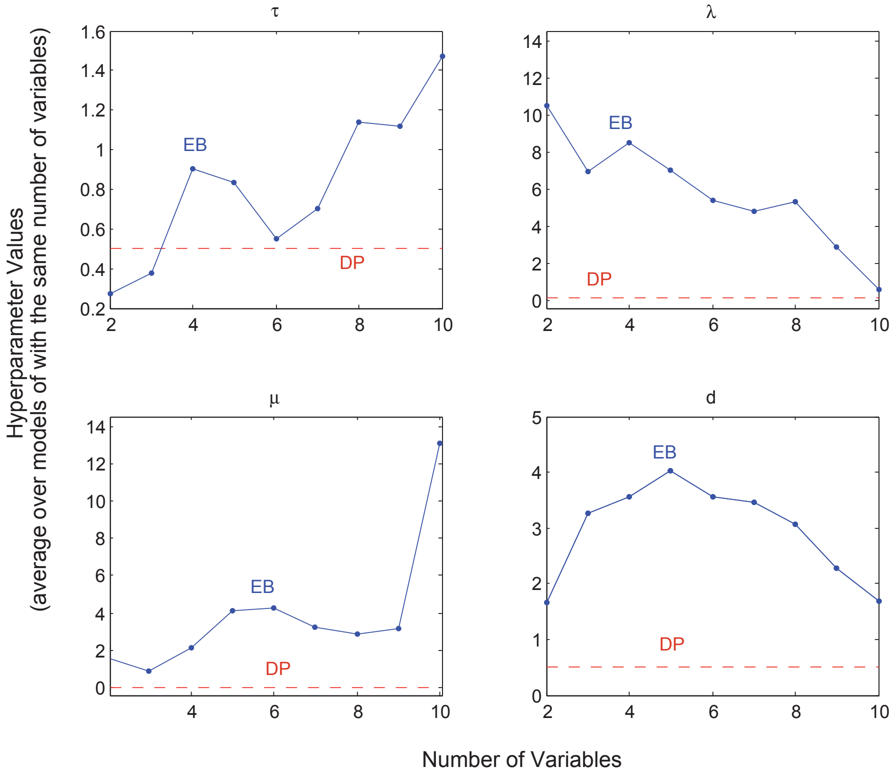
Although one could argue that an optimal strategy is to use tight priors with strong GDP inertia during expansions and to employ diffuse priors during recessions, when all economists agree that uncertainty is larger, it is not straightforward to know in real-time when it is the right moment to switch. Therefore, we compare the empirical bayes approach described above with the alternative of

[10]This is quite often the case because expansionary periods are long and stable, while recessions are short.

setting very diffuse priors for all models independently of their size. The values chosen for the diffuse priors are given in Figure 1.

The main advantage of this approach lies in its simplicity. When the number of variables becomes moderately large, setting up informative priors for all possible models using the Empirical Bayes method could take years[11].

Figure 1: The Tightness of the Prior



The figures display the average value of the hyperparameters estimated with the EB approach for models with the same number of predictor variables. The number of models of size equal to 2, 3, 4, 5, 6, 7, 8, 9 and 10 is equal to 10, 45, 120, 210, 252, 210, 120, 45 and 10, respectively. The precise definition of each one of the hyperparameters can be found in the appendix. τ: overall tightness of the prior, λ: one-unit-root prior (co-persistency prior), μ: no-cointegration prior (own persistency prior), d : rate of decay for the prior shrinking the lags.

[11]On average, optimizing the hyperparameters to maximize forecast accuracy over the training sample takes on average one minute with a 2.20GHz processor. This means that we can construct priors for 1,023 models (resulting from all combinations of GDP with 10 predictor variables) in 17 hours. Obtaining priors for 1,048.576 (resulting from all combinations of GDP with 20 predictor variables) is unfortunately not feasible, since it would take roughly 2 years.

3.3 Information Subsets

All the projections (see equation 1) are conditional on information *sub*sets available approximately one and a half months before the statistical agency publishes its official release. Justifying the use of a particular forecasting model and the selection of conditioning information is a challenging task. Researchers and analysts are often satisfied with a model that yields accurate forecasts for a given sample period. In this paper, however, we will consider all the linear projections one can construct with all possible combinations of GDP and the indicators contained in two different information sets.

$\mathbf{\Omega^1}$: The first information set contains the 11 key variables shaded in Figure 2. Those indicators provide leading information about the GDP components and the aggregate business cycle behaviour of the economy. This information set includes eight of the variables selected by Camacho and Perez-Quiros (2010b): total employment, retail trade confidence indicator, services PMI, industrial confidence, industrial production, sales of big firms, real exports and imports. In addition, we incorporate indicators that are highly correlated with the aggregate GDP growth time series: the economic sentiment indicator, which tracks very closely the year-on-year GDP growth figures, and the stock exchange index (IBEX'35), which is related with nominal long-term growth of the economy.

$\mathbf{\Omega^2}$: The second information set extends the first one by including additional indicators for some of the GDP subcomponents ($\Omega^1 \subset \Omega^2$). This set includes car registrations, air transportation, building permits, hotel stays, construction employment, industry PMI, the consumer confidence indicator, total sales and imported oil price in euros. Although one could argue that the first subset is sufficiently representative of the Spanish business cycle, our aim is to understand whether further accuracy gains can be achieved by enlarging the size of the models.

4 Empirical Results

Table 4 summarizes the basic ex-ante forecasting strategies that we evaluate. With information set Ω^1, the projection equation 1 will allow us to construct a total of $1,023$ models with N ranging from 2 to 10. The larger information set Ω^2 will allow us to construct a total of $262,144$ models with N ranging from 10 to 19.

Table 4: Strategies for GDP growth NOWCASTING

Information & Model Set	**Prior Elicitation**	**Evaluation Sample**	
Small Information Set Ω^1 (1023 Models of size 2-10)	Empirical Bayes (EB)	-	2008Q4-2010Q2
	Diffuse Priors (DP)	2006Q3-2008Q3	2008Q4-2010Q2
Extended Information Set Ω^2 (262144 Models of size 11-20)	Diffuse Priors(DP)	2006Q3-2008Q3	2008Q3-2010Q2

Comparing both EB and DP strategies for the estimation of all models included in the small information set Ω^1 will shed light on the usefulness of *ex-ante* prior information as a way to improve forecast accuracy only over the second subsample. An alternative option to achieve forecast accuracy is to benefit from a larger information set, Ω^2. We will explore the possibility that the larger information set under the DP strategy provides accuracy gains beyond those given by the use of Ω^1 under the same prior elicitation strategy. This evaluation can be conducted on the basis of both subsamples, since it does not require any training period to select priors or combination weights.

4.1 Gains from the Empirical Bayes Approach

In this subsection, we aim to provide evidence about the advantages of the Empirical Bayes method (EB) over the use of diffuse priors (DP). We will analyze the forecasts based on the $1,023$ different models that can be constructed with Ω^1. All projections are obtained with the information available approximately one and a half months before the statistical agency publishes the national accounts.

A simple analysis of the root mean squared errors in Tables 5 and 6 reveals that the average forecast under EB reduces the RMSE compared with the DP strategy by more than 10% throughout the second subsample[12]. Figures 2 and 3 provide visual evidence going beyond the summary statistics discussed above. These figures also display the forecasting distribution of the 10% top-performing models (fanchart) over the training sample in addition to the simple mean of all models (dashed line). Figure 2 reveals that prior elicitation based on the training sample helps to achieve excellent forecasts during the 2008Q4-2010Q2 period with a weighted average of the top 20 models (solid line). However, the preference for using either the 20 best (ex-ante) forecasting models over a weighted average of the whole set of models is only easily justified ex-post.

Nevertheless, the gains from the EB approach with respect to the DP strategy are also visible in Figures 4 and 5, which show root mean squared errors of increasingly large forecast combinations for the evaluation period. These figures show that the *combination* of models is always more accurate when the EB method is used, independently of the number of models used to construct the combined forecast. The results, however, do not seem to be statistically significant in the light of Figure 7. All the projection models obtained with Ω_1 under either the DP or the EP approach yield thousands of time series of forecast errors corresponding to our evaluation sample (2008Q4-2010Q2). These graphs repre-

[12]This result holds regardless of whether the forecast error is computed on the basis of the "preliminary" (Table 5) or the "final" GDP release (Table 6).

sent the probability distributions of all these forecast errors, which is very similar regardless of the prior elicitation strategy.

Table 5: Forecast accuracy with respect to the "preliminary" releases

Information & Model Set	**Prior Elicitation**	**RMSE for the simple average**	
		2006Q3-2008Q3	2008Q4-2010Q2
Small Information Set Ω^1 (1023 Models of size 2-10)	Empirical Bayes (EB)	-	0.308
	Diffuse Priors (DP)	0.236	0.358
Extended Information Set Ω^2 (262144 Models of size 11-20)	Diffuse Priors(DP)	0.339	0.143

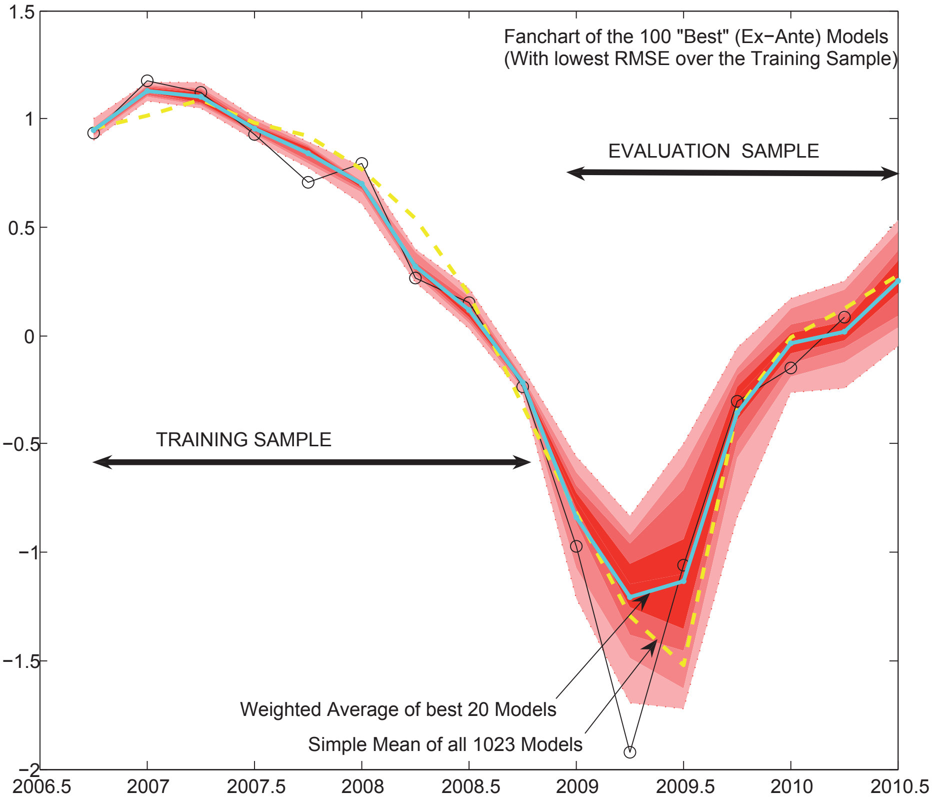
Comparing both EP and DP strategies for the estimation of all models included in the small information set Ω^1 sheds light on the usefulness of *ex-ante* prior information as a way to improve forecast accuracy over the second subsample. The results show that the DP strategy yields a RMSE 16% larger than the EB approach when the errors are computed on the basis of the first available GDP growth rates. An alternative option to achieve forecast accuracy is to benefit from a larger information set, Ω^2. We also explore the possibility that the larger information set under the DP strategy provides accuracy gains beyond those given by the use of Ω^1 under the same prior elicitation strategy. Under the DP strategy, the larger information set Ω^2 allows us to achieve much higher forecast accuracy than that resulting from Ω^1.

Table 6: Forecast accuracy with respect to the "last available" releases

Information & Model Set	**Prior Elicitation**	**RMSE for the simple average**	
		2006Q3-2008Q3	2008Q4-2010Q2
Small Information Set Ω^1 (1023 Models of size 2-10)	Empirical Bayes (EB)	-	0.246
	Diffuse Priors (DP)	0.251	0.295
Extended Information Set Ω^2 (262144 Models of size 11-20)	Diffuse Priors(DP)	0.366	0.167

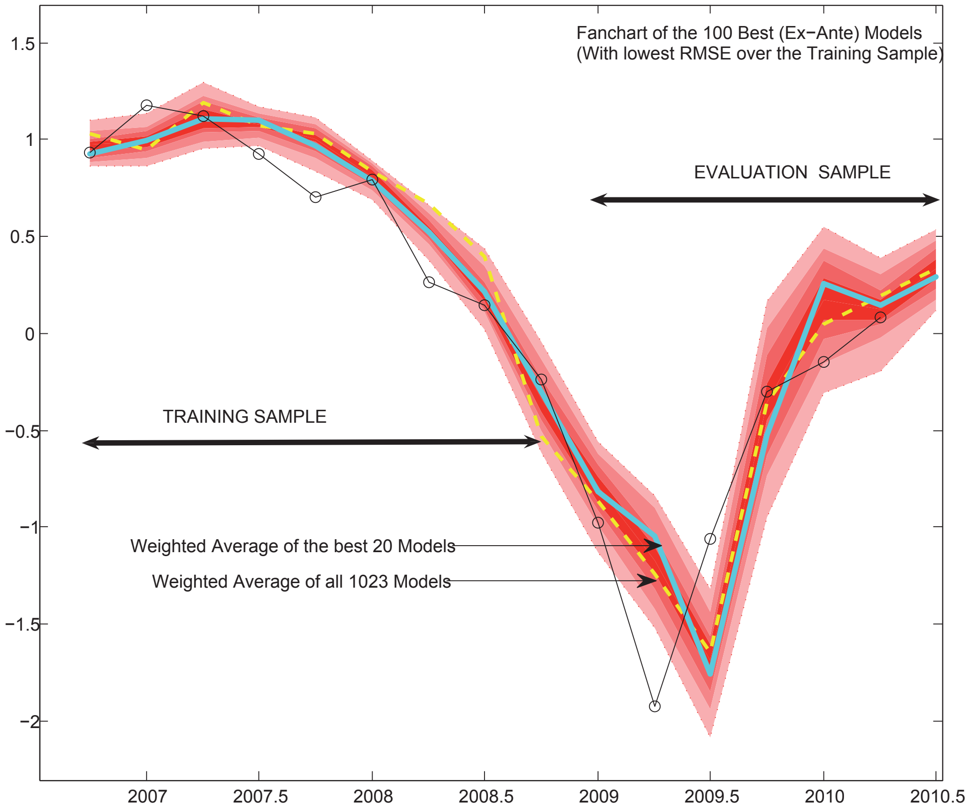
Comparing both EP and DP strategies for the estimation of all models included in the small information set Ω^1 sheds light on the usefulness of *ex-ante* prior information as a way to improve forecast accuracy over the second subsample. The results show that the DP strategy yields a RMSE 20% larger than the EB approach when the errors are computed on the basis of the last available vintage for GDP growth. An alternative option to achieve forecast accuracy is to benefit from a larger information set, Ω^2. Under the DP strategy, the larger information set Ω^2 allows us to achieve much higher forecast accuracy than that resulting from Ω^1.

Figure 2: Nowcasts conditional on Ω^1 (Empirical Bayes)



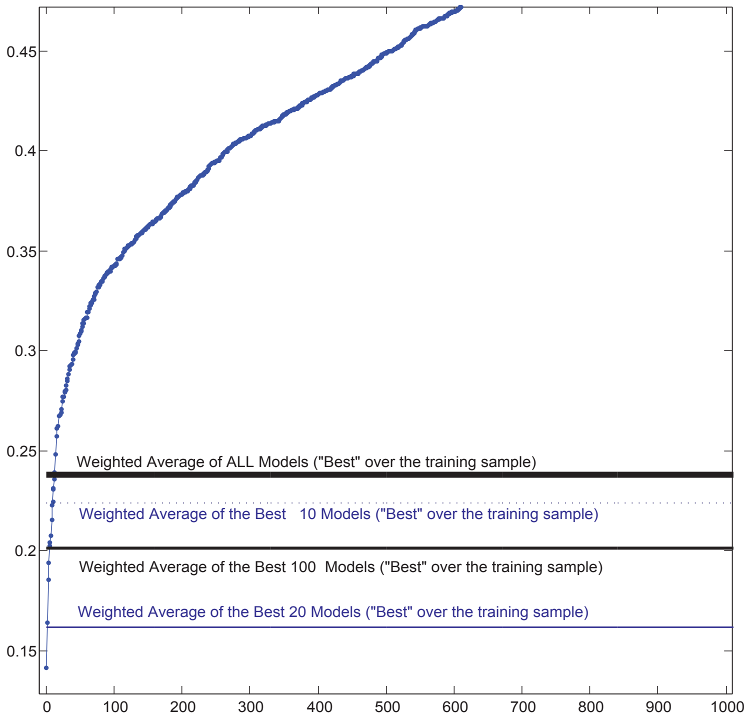
The black circles represent real GDP growth as initially published by the statistical agency. Given that we use the training sample to form priors, it is not surprising that the 10% best performing models provide a perfect fit for GDP growth. The question of interest is whether those models "selected" on the basis of their performance are able to continue being accurate over the evaluation sample.

Figure 3: Nowcasts conditional on Ω^1 (Diffuse Prior)



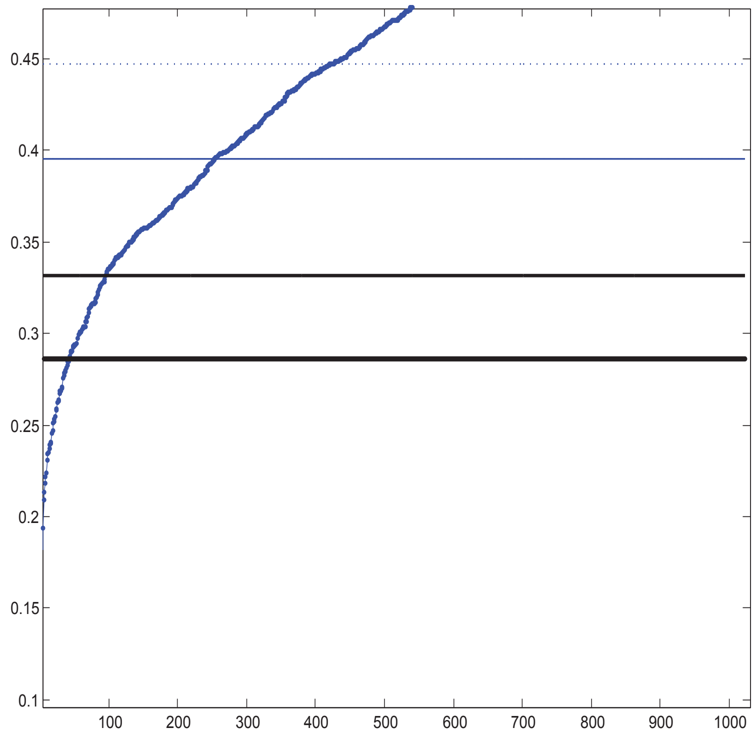
The black circles represent real GDP growth as initially published by the statistical agency. The thick line represents the weighted average nowcast of the 20 models with smallest RMSE over the first subsample. The question of interest is whether those models "selected" on the basis of their performance over the training sample are able to continue being accurate over the evaluation sample. Alternatively, the dashed line is a simple average of all 1023 models. Since this strategy does not require any prior information from the first subsample, it can be evaluated over the whole recession episode (not only over the so-called evaluation sample).

Figure 4: RMSE 2008Q4-2010Q2, (Empirical Bayes, Ω^1)



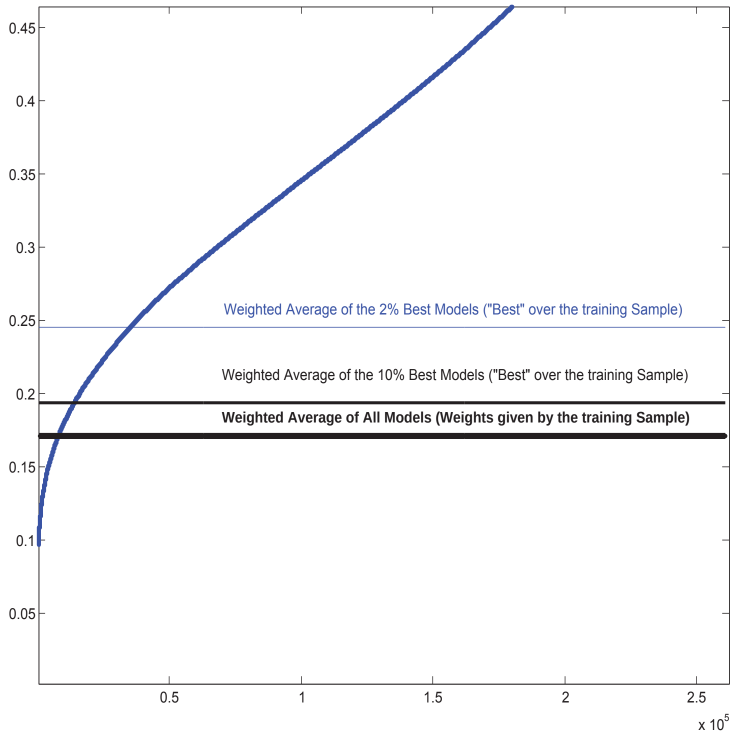
The Root Mean Squared Error (RMSE) for each model is computed on the basis of real-time out-of-sample forecast errors for GDP growth. The prediction error is defined as the difference between the nowcast and the last available GDP growth release published by the statistical agency. The RMSEs of all models are sorted in ascending order. The dotted line corresponds to the RMSE associated to the weighted average of the best 10 performing models over the training sample. Averaging over the top 20 results on a very large increase in forecast accuracy. Actually, the figure shows that there is only one model with better forecast accuracy (one point below the thinnest solid line). Finally, incorporating all models does not help to achieve a further reduction in RMSE. Here, the training sample 2006Q3-2008Q3 is used for both forming the priors and choosing the forecast combination weights.

Figure 5: RMSE 2008Q4-2010Q2, (Diffuse Prior, Ω^1)



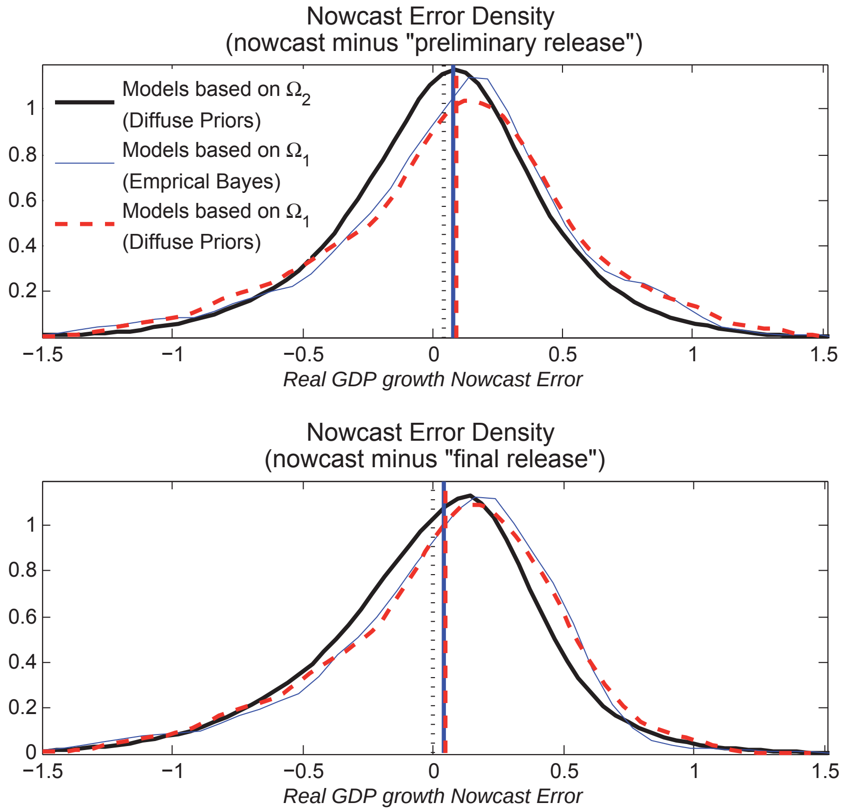
The Root Mean Squared Error (RMSE) for each model is computed on the basis of real-time out-of-sample forecast errors for GDP growth. The prediction error is defined as the difference between the nowcast and the last available GDP growth release published by the statistical agency. The RMSEs of all models are sorted in ascending order. The dotted line corresponds to the RMSE associated to the weighted average of the best 10 performing models over the training sample. Averaging over 20 and 100 models increases forecast accuracy. The thickest line is associated to the weighted average of all models. Here, the training sample 2006Q3-2008Q3 is used only to choose the forecast combination weights.

Figure 6: RMSE 2008Q4-2010Q2, (Diffuse Prior, Ω^2)



The Root Mean Squared Error (RMSE) for each model is computed on the basis of real-time out-of-sample forecast errors for GDP growth. The prediction error is defined as the difference between the nowcast and the last available GDP growth release published by the statistical agency. The RMSEs of all models are sorted in ascending order. The dotted line corresponds to the RMSE associated to the weighted average of the best 2% performing models over the training sample. When all models are considered in the weighted average, i.e. the thickest line, forecast accuracy increases (RMSE goes down). It can be shown that a simple average, i.e. giving the same *weight* to all models would yield exactly the same value.

Figure 7: Density of Forecast Errors resulting from Ω^1 (DP vs EB) and Ω^2 (DP)



All the projection models obtained under Ω_1 and Ω_2 yield thousands of time series of forecast errors corresponding to our evaluation sample (2008Q4-2010Q2). These graphs represent the probability distributions of all these forecast errors. The upper figure shows that when the small information set (Ω_1) is used, both EB and DP strategies yield a very similar *nowcast* error density with mean slightly larger than zero, which is consistent with a slight overprediction of GDP growth over the most severe part of the recession. When Ω_2 is used, the *nowcast* error density shifts towards the left and concentrates more probability mass around zero. Note that the mean of the distributions, which is marked with vertical lines, does not necesarily coincide with the mode.

4.2 Gains from a Larger Information set

The previous subsection described the gains derived from exploiting pre-sample information to elicit priors. In this section, we present an alternative strategy for achieving information gains. Rather than modifying our priors, we enlarge the number of predictor variables in the hope of improving forecast accuracy. When the set of candidate variables expands, the total number of models that can be constructed increases exponentially. Whereas the use of Ω^1 has allowed us to combine a maximum of 10 predictor variables with GDP, the larger information set, Ω^2 , allows us to exploit the information from a total of 19 indicators.

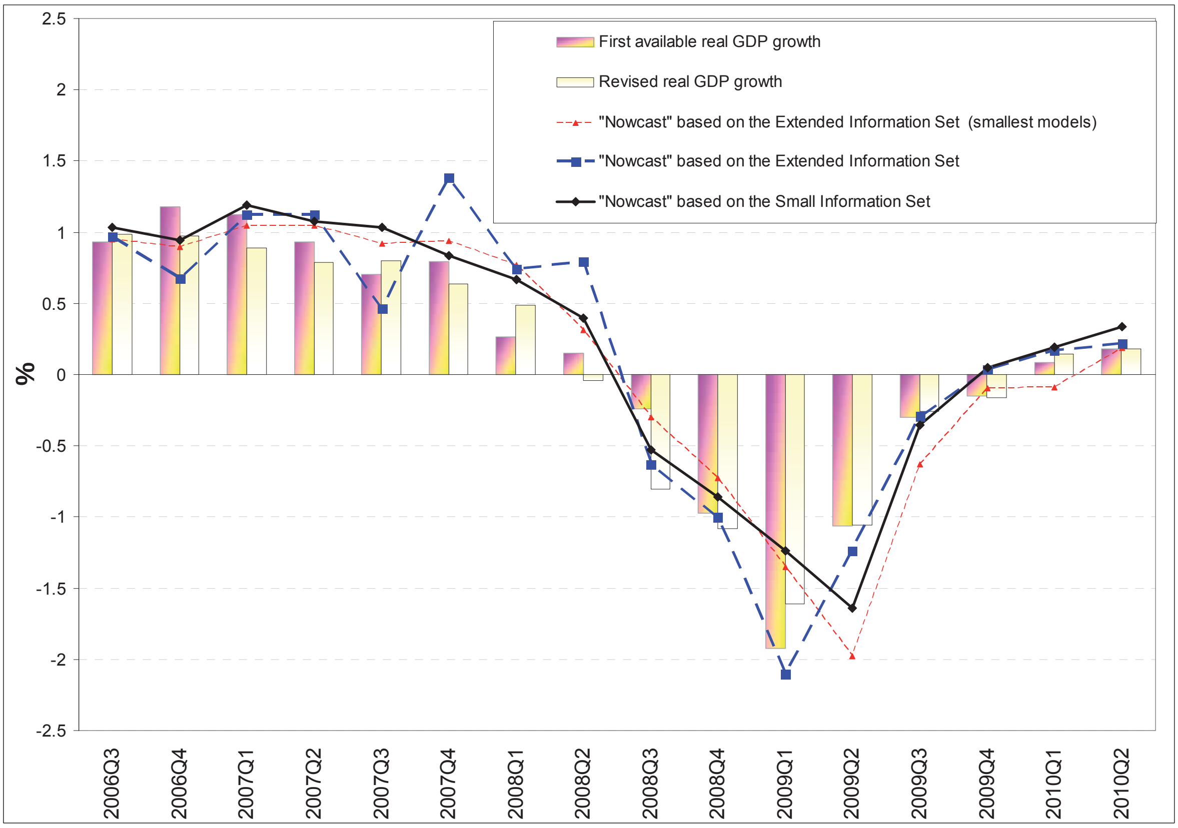
The presence of collinearity in the data could lead us to think that the 10 predictor variables of Ω^1 are sufficiently representative, and enlarging the information set is redundant. However, the larger information set Ω_2 allows us to aggregate forecasts coming from larger models. In particular, we propose a combination of medium-sized models that incorporates a number of indicators ranging from ten to nineteen. We expect that larger models are more likely to identify the multiple factors underlying business cycle fluctuations, thereby decreasing the risk of model misspecification and improving forecast accuracy. Figure 8 clearly shows that the alternative model combination option that incorporates only two or three indicator variables belonging to the large information set Ω_2 produces forecasts that are highly correlated with those obtained with the small set Ω_1. Thus, the gains of using a larger information set come from the ability to use larger models.

The RMSE results are interesting when we compare the two subsamples in which we divide the recession. Both Tables 5 and 6 provide overwhelming evidence in favour of Ω^2 for the second subsample, 2008Q4-2010Q2, which is visualized in Figure 11. However, the gradual slowdown registered over the 2006Q3-2008Q3 period has been predicted more accurately with the reduced information set Ω^1, as shown in Figure 8. Notice that a fair comparison over the first subsample should be based on the mean forecast and not on the weighted averages. After a very successfull projection for 2007Q3, larger models tend to over-react to the news regarding 2007Q4 and subsequent quarters. As a result, RMSEs based on Ω^2 deteriorate for the first subsample.

To understand the added value of Ω^2 over the second subsample, we can analyze the density of the realized real-time forecast errors over the period 2008Q4-2010Q2. Because we look at thousands of models, our results are unlikely to be driven by data snooping[13]. The top panel of Figure 7 refers to the forecast errors based on the preliminary real GDP growth figures, whereas the graph at the bottom refers to the so-called final release, or lastest available data. We can observe that Ω^2 yields a forecast error density that is more centered on zero. This is an intuitive way of suggesting that the limited information set Ω^1 may result in poor forecasting performance throughout the most severe phase of the recession. However, the differences are smaller when the errors are defined in terms of the final release (bottom panel).

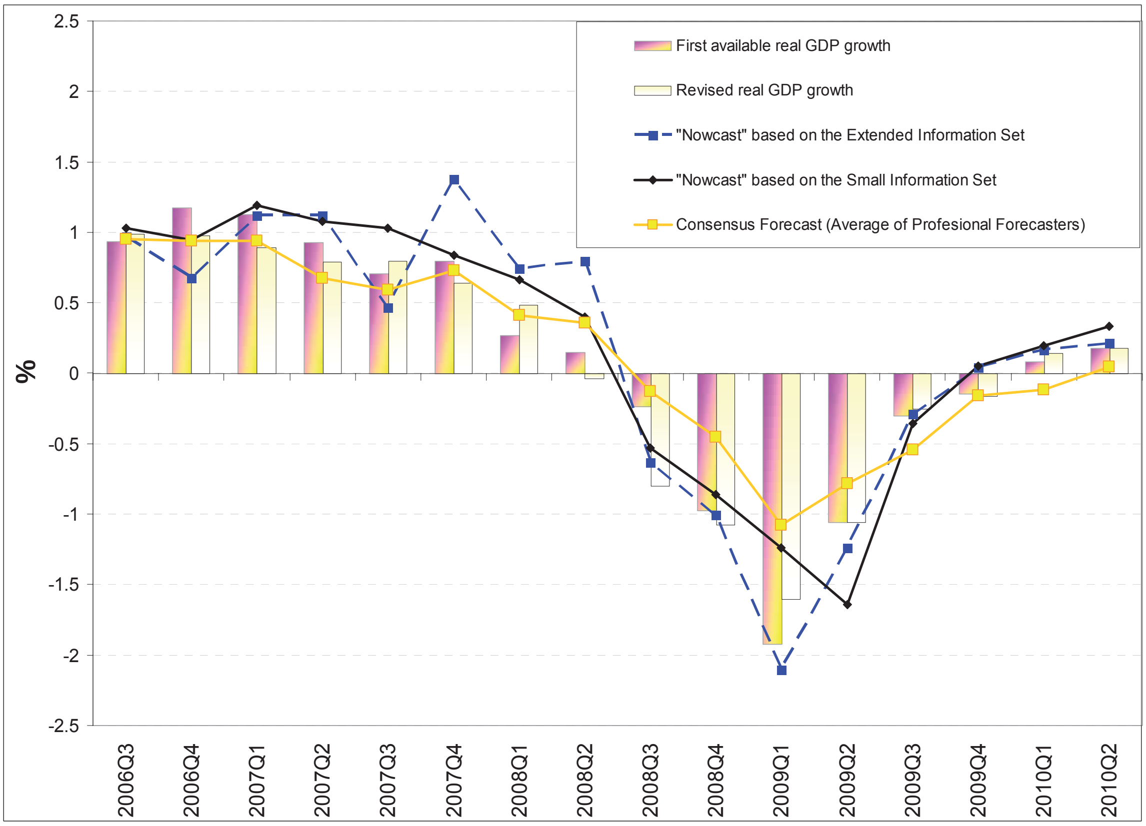
[13]The use of a "single" model to evaluate the gains derived from the Empirical Bayes Approach proposed in this paper would be misleading. Given the small size of our evaluation sample, which corresponds to the current recession episode, the results could be overly dependent on the choice of the conditioning information set. That is, one could *randomly* choose a model for which the Empirical Bayes (EB) approach happens to yield significantly better GDP forecasts than the use of less informative priors (DP) and wrongly conclude that prior elicitation according to the first method is superior.

Figure 8: *Small* models based on the extended information set Ω_2



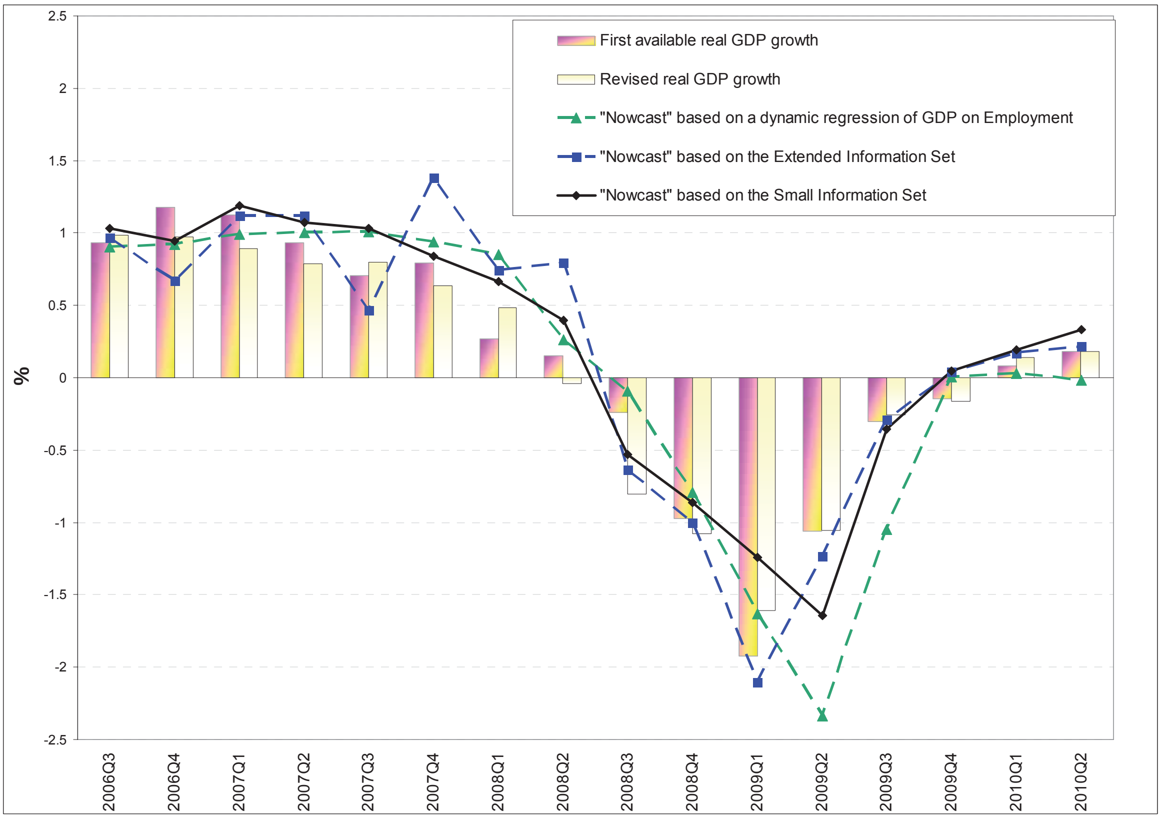
The figure compares the nowcasting performance of a simple average of *large* models based on the set Ω_2 (dashed line with squares) with the one based on the smaller information set Ω_1, which only contains 10 economic indicators other than GDP (solid line). Moreover, we show that a combination of all (*small*) models one can construct by combining two and three indicators available in the information set Ω_2 does not yield accurate nowcasts during the most severe part of the recession (2008Q3-20010Q2). Thus, nowcast combinations based on Ω_2 are successful only when a medium or large number of variables is incorporated in the individual forecasting equations.

Figure 9: Comparison with the "Consensus Forecast"



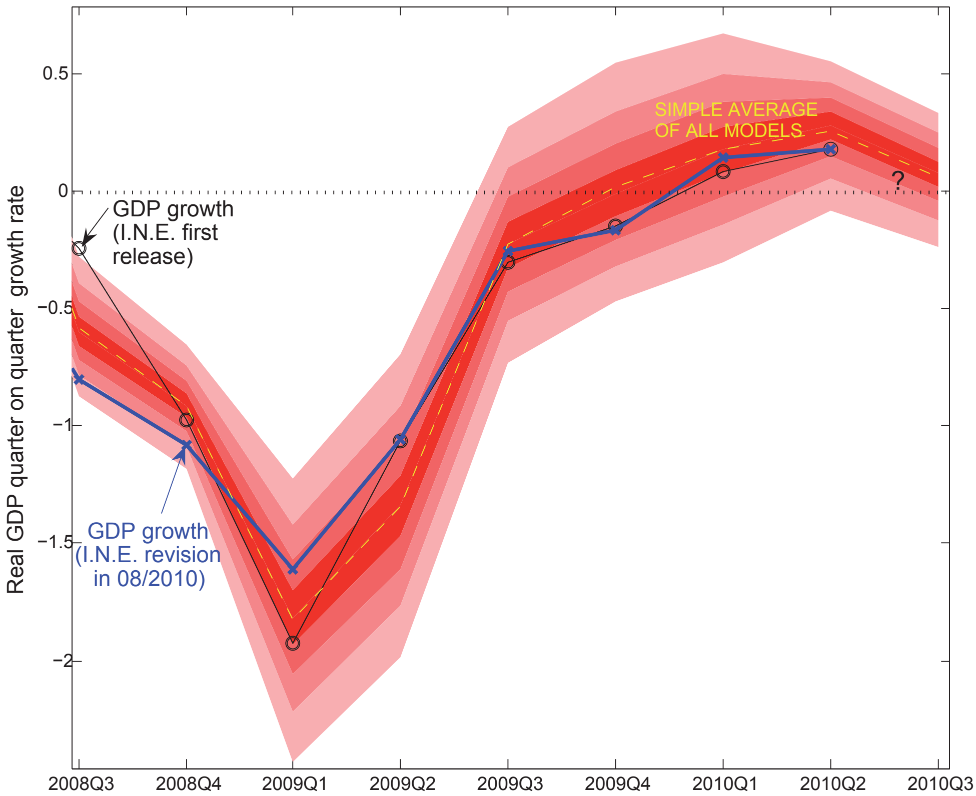
The figure illustrates the forecasting ability of the mean of the survey of profesional forecasters compiled by Consensus Economics and published in their montly publication "Consensus Forecast". This comparison is quite meaningful, since it is also an aggregation of individual forecasts. Moreover, we have selected only the publications of the months January, April, July, and October, which coincide with our nowcasting calendar. In addition to that, it is worth enphasizing that since *Consensus Forecasts* typically refer to year-on-year growth rates, it is necessary to use a real-time database in order to recover quarter-on-quarter growth, which is our measure of interest.

Figure 10: Employment as a predictor variable



The figure illustrates the excellent forecasting ability of a simple GDP projection on employment during of 2008Q2-2009Q1 period and compares it with two of our forecasting strategies. The projections represented with triangles result from a simple regression of GDP on its first lags and current and past values of employment. The disadvantage of this model is that when employment figures start to improve during the second quarter of 2009, the turning point in growth that occurs in that period cannot be predicted. More generally, we have found that the bad performance of small models around this turning point is typically driven by a strong contribution of the GDP inertia.

Figure 11: Nowcasts conditional on Ω^2 (Diffuse Prior)



The black circles represent real GDP growth as initially published by the statistical agency. The fanchart represents a 90% forecasting interval that takes into account model uncertainty. The dashed line is a simple average of all 262144 models. This graph also represents the projection exercise for 2010Q3, which has been conducted at the beginning of October.

4.3 How Accurate Are Our Nowcasts?

Spanish real GDP growth is a very smooth time series with significantly smaller variance than growth figures published in other countries like the US or Germany. This is due both to economic reasons, e.g. the stabilizing effect of imports, and to measurement issues concerning the procedures used by the statistical agency to estimate an efficient signal of Spanish economic growth. As a consequence, Spanish GDP growth is highly predictable. This implies that most of the forecasting methods have serious difficulty in improving on the forecast accuracy of a random walk model for the growth rates, which is a very common benchmark in macroeconomic forecasting applications.

Because our main nowcasting strategies are based on model combinations, it is interesting to compare their performance with the mean prediction resulting from the survey of professional forecasters compiled by *Consensus Economics* and published in their monthly magazine "Consensus Forecast". Figure 9 shows that the *Consensus Forecast* follows GDP growth very closely until 2008Q2, where it fails to predict the first negative quarterly growth figure. Both of our forecast combination strategies (Ω_1 and Ω_2, with diffuse priors) and the statistical agency itself, in its initial announcement, were unable to predict the negative growth rate in 2008Q2. However, the large decline in economic activity registered over the subsequent quarter is perfectly predictable by our forecast combinations and slightly underestimated by the *Consensus Forecast*. Finally, the growth for the three subsequent quarters is clearly over-predicted by the *Consensus Forecast*. This example illustrates the difficulty of the forecasting practice over the most severe phase of the recession.

Relative Forecast Accuracy of the Forecast Combinations

Table 7 provides the RMSE of the different forecasting schemes divided by the RMSE of the random walk forecast. The reputation of professional forecasters is generally based on their ability to forecast the preliminary or first available releases. As seen in the left panel of the table, the Consensus Forecast provides the highest forecast accuracy over the first subsample, which corresponds to the gradual start of the deceleration phase. However, when the whole sample is considered, the Consensus Forecast is less precise, regardless of whether our focus of interest is the preliminary or the final GDP growth release. The most significant result is the excellent forecast accuracy achieved over the second subsample by combining projections conditional on subsets of Ω_2, the so-called *Extended Information Set*.

Table 7 also compares our forecast combination strategies with the use of a single model. Not surprisingly, the autoregressive distributed lag model that incorporates all the indicators included in Ω_2 results in a very low RMSE over the second subsample, although it is outperformed by the simple forecast combination.

Sensitivity Analysis

The success of our model-based forecast combinations may have a very simple explanation. Among all the time series included in any given information set, it is unavoidable to find measurement errors in the form of outliers or seasonal effects that are not always easy to correct in real-time[14]. Bad quality data may contribute to deteriorate forecast accuracy. By combining the projection models obtained with the noisy data with those that do not contain it, the negative impact of the noisy data is reduced. It is precisely the presence of measurement errors an important motivation for the use of factor models, since they work as a filter that extracts a clean estimate of the business cycle factors without

[14]Although we use TRAMO-SEATS as an automatic way of adjusting the series in real time, it is impossible to reproduce the judgemental adjustments of sectoral experts or the adjustments to the series made by the statistical agency itself.

the need to discard noisy data. In our case, given the large amount of synchronization observed over the business cycle, discarding one or few noisy indicators from the conditioning information set is unlikely to entail misspecification problems. Thus, the success of our forecast combination strategy is based on the same principles as the success of the factor models: a) a considerable amount of comovements/collinearity, which they capture with pervasive *common* factors, and b) the presence of measurement errors, or *idiosyncratic* components.

Table 8 illustrates this idea by describing the forecast accuracy of our model combination strategy based on Ω_2 when each one of the predictor variables is ignored one at a time. Independently of whether we use preliminary data (left panel) or revised data (right panel) to compute our relative RMSE measure of fit, none of the *exclusions* results in a significant deterioration of forecast accuracy for the whole sample, as expected. Conversely, when we focus on the right-hand side of the table, we can observe that the RMSE over the first subsample improves considerably when either building permits or the retail trade confidence indicator is excluded from the forecast combination. When both of them are excluded (see the last row of the first section of Table 8), the relative RMSE decreases to such an extent that our nowcasts can be considered to be even more precise than the first release of the statistical agency itself. This is the conclusion one can draw by comparing these results with the RMSE associated with the first release when we think of it as a forecast of the latest available data (see last row of the table).

Table 7: RMSE of alternative nowcasting procedures *divided by* ***RMSE of a random walk forecast***

	Forecasting the **preliminary release**: full recession 2006Q3-2010Q2	Forecasting the **preliminary release**: start of deceleration phase 2006Q3-2008Q3	Forecasting the **preliminary release**: around the turning point in growth rates 2008Q4-2010Q2	Forecasting the **last available release**: full recession 2006Q3-2010Q2	Forecasting the **last available release**: start of deceleration phase 2006Q3-2008Q3	Forecasting the **last available release**: around the turning point in growth rates 2008Q4-2010Q2
SIMPLE MODEL COMBINATIONS (equal weights)						
Small Information Set (DP) *Introducing all variables included in Ω_1 (1023 models with 2-11 variables)*	**0.64**	0.91	0.56	**0.55**	0.66	0.48
Extended Information Set (DP) *Introducing all variables included in Ω_2 (262144 models with 11-20 variables)*	**0.58**	1.31	0.22	**0.60**	0.96	0.27
COMBINATION OF PROFESIONAL FORECSTERS (equal weights)						
Consensus Forecast	**0.64**	0.64	0.64	**0.64**	0.72	0.59
A FEW SELECTED MODELS						
THE LARGEST MODEL (20 variables in Ω_2)	**0.73**	1.68	0.25	**0.78**	1.22	0.41
MEDIUM SIZED MODEL (11 variables in Ω_1)	**1.02**	1.44	0.91	**0.83**	0.89	0.79
NAIVE MODEL (2 variables: GDP and Employment)	**0.92**	0.98	0.91	**0.92**	0.84	0.96

Table 8: Sensitivity Analysis (RMSE *divided by* ***RMSE of a random walk forecast***)

	Forecasting the **preliminary release**			Forecasting the **last available release**		
	full recession	start of deceleration phase	around the turning point in growth rates	full recession	start of deceleration phase	around the turning point in growth rates
	2006Q3-2010Q2	2006Q3-2008Q3	2008Q4-2010Q2	2006Q3-2010Q2	2006Q3-2008Q3	2008Q4-2010Q2
SIMPLE COMBINATIONS (equal weights)						
Extended Information Set (DP) *Introducing all variables included in* Ω_2	**0.58**	1.31	0.22	**0.60**	0.96	0.27
Excluding Industrial Confidence Indicator	**0.57**	1.27	0.22	**0.58**	0.93	0.27
Excluding Retail Trade Confidence Indicator	**0.56**	1.29	0.17	**0.54**	0.87	0.22
Excluding PMI Services	**0.63**	1.43	0.23	**0.63**	1.00	0.31
Excluding Total Employment	**0.59**	1.35	0.21	**0.61**	0.98	0.27
Excluding Car Registrations	**0.61**	1.32	0.29	**0.61**	0.98	0.26
Excluding Construction Employment	**0.61**	1.38	0.21	**0.62**	1.01	0.26
Excluding Consumer Confidence Indicator	**0.58**	1.34	0.19	**0.58**	0.97	0.20
Excluding Economic Sentiment Indicator	**0.62**	1.39	0.22	**0.67**	1.09	0.26
Excluding PMI Industry	**0.60**	1.34	0.25	**0.61**	0.96	0.31
Excluding IBEX'35 (Stock Exchange Index)	**0.59**	1.34	0.22	**0.62**	0.99	0.28
Excluding Industrial Production Index (non-energy)	**0.63**	1.30	0.35	**0.62**	0.94	0.37
Excluding Sales (big firms)	**0.56**	1.26	0.22	**0.56**	0.90	0.27
Excluding Hotel Stays by foreigners	**0.60**	1.34	0.24	**0.62**	0.97	0.33
Excluding Sales (non-financial)	**0.62**	1.41	0.23	**0.70**	1.10	0.35
Excluding Imported Oil Price in Euros	**0.67**	1.35	0.40	**0.66**	0.99	0.40
Excluding Real Exports	**0.66**	1.34	0.38	**0.65**	0.99	0.38
Excluding Real Imports	**0.61**	1.29	0.31	**0.65**	0.98	0.39
Excluding Air Transportation. (Metric Tones)	**0.58**	1.31	0.23	**0.60**	0.95	0.28
Excluding Building Permits	**0.61**	1.26	0.34	**0.51**	0.75	0.32
Excluding Building Permits and Retail Trade Confidence indicator	**0.56**	1.21	0.28	**0.38**	0.62	0.16
Average of PROFESIONAL FORECASTERS (Consensus Economics)	0.64	0.64	0.64	**0.64**	0.72	0.59
First Release of the I.N.E. viewed as a forecast of the last available vintage	- -	-	-	**0.41**	0.64	0.22

5 Conclusion

The gradual slowdown in economic activity that took place during 2007 and the subsequent recession and recovery provide an adequate environment to test our forecasting models. After all, it is precisely in these periods of great uncertainty when analysts and policy-makers want to have accurate forecasts.

This paper provides evidence about the predictability of Spanish GDP growth one and a half months before the official figures are published by the statistical agency. We show that Bayesian dynamic regressions, which allow us to obtain projections conditional on subsets of available predictor variables, yield accurate forecasts in real time. Overall, our nowcasts are more accurate than the mean prediction resulting from the survey of professional forecasters published by "Consensus Forecast".

To our knowledge, our paper presents the first real-time "nowcasting" exercise with medium sized autoregressive distributed lag models. In general, the larger the number of indicators included in a regression, the smaller the risk of model misspecification. This requires the estimation of a very large number of parameters, which could lead to in-sample overfitting and large out-of-sample forecast errors. The potential multicolinearity problems arrising from the large amount of sincronization among the predictor variables is offset by the use of priors or "inexact" restrictions originated in the VAR literature.

As shown by De Mol et al. (2008), forecasts based on large bayesian regressions can be highly correlated with those resulting from static principal components. Thus, it makes sense to think that large dynamic regressions may be able to capture the business cycle co-movements without the need to impose an analytical dynamic factor structure. The large and medium sized bayesian VARs developed by Banbura et al. (2010a) to forecast monthly US macro variables illustrate this idea and motivate the use of dynamic regressions also for the "nowcasting" practice.

References

[1] Armesto M.T., K.M. Engenmann and M.T. Owyang (2010) . "Forecasting with Mixed Frequencies". *Federal Reserve Bank of St. Louis Review*, *92*, 521-536.

[2] Andersson, M.K., and S. Karlsson (2008). "Bayesian forecast combination for VAR models". In: S. Chib, W. Griffiths, G. Koop and D. Terrel (Eds), *Bayesian Econometrics*, 501-524.

[3] Angelini, E., G. Camba-Mendez, D. Giannone, G. Runstler, and L. Reichlin (2010). "Short-Term Forecasts of Euro Area GDP Growth". *The Econometrics Journal*, *forthcoming.*

[4] Banbura M., and M. Modugno (2010). "Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data". *European Central Bank Working Paper Series*, 1189.

[5] Banbura M., D. Giannone and L. Reichlin (2010a). "Large Bayesian VARs". *Journal of Applied Econometrics*, **25**, 71-92.

[6] Banbura M., D. Giannone and L. Reichlin (2010b). "Nowcasting". *Oxford Handbook on economic Forecasting*, in Michael P. Clements and David F. Hendy, *forthcoming.*

[7] Banbura M., and G. Runstler (2010). "A look into the factor model black box. Publication lags and the role of hard and soft data in forecasting GDP". *International Journal of Forecasting*, forthcoming.

[8] Bates J.M. and C.W.J. Granger (1969). "The Combination of Forecasts". *Operations Research Aquarterly*, **20**, 451-468.

[9] Clark T. and M.W. McCracken (2010). "Averaging Forecasts from VARs with Uncertain Instabilities". *Journal of Applied Econometrics*, **25**, 5-29.

[10] Cuevas, A., E. Quillis (2010). "A Factor Analysis for the Spanish Economy (FASE)". Mimeo

[11] Camacho M., and R. Domenech (2010). "MICA-BBVA: A factor model of economic and financial indicators for short-term GDP forecasting". *BBVA Research Department Working Paper* , 1026.

[12] Camacho M., G. Perez Quiros (2010a). "Introducing the Euro-Sting". *Journal of Applied Econometrics*, **25**, 663-694.

[13] Camacho M., G. Perez Quiros (2010b). "Ñ-sting: España Short Term Indicator of Growth". *The Manchester School*, forthcoming.

[14] Croushore, D. and T. Stark (2001). "A Real-Time Data Set for Macroeconomists". *Journal of Econometrics*, **105**, 111-130.

[15] Croushore, D. and T. Stark (2002). "Forecasting with a Real-Time Data Set for Macroeconomists". *Journal of Macroeconomics*, **24**, 507-531.

[16] Chow G.C., A. Lin (1971). "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series". *The Review of Economics and Statistics*, **53**, 372-375.

[17] Clemen, R. T. (1989). "Combining forecasts: A review and annotated bibliography". *International Journal of Forecasting*, **5**, 559-583.

[18] De Mol C., D. Giannone and L. Reichlin (2008). "Forecasting Using a Large Number of Predictors: Is Bayesian Regression a Valid Alternative to Principal Components?". *Journal of Econometrics*, **146**, 318-328.

[19] Diebold, F.X., and P. Pauly (1987). "Structural Change and the Combination of Forecasts". *Journal of Forecasting*, **6**, 21-40.

[20] Doan, T., R. Litterman and C.A. Sims (1984). "Forecasting and Conditional Projection Using Realistic Prior Distributions". *Econometric Reviews*, **3**, 1-100.

[21] Eklund J. and S. Karlsson (2007). "Forecast Combination and Model Averaging Using Predictive Measures". *Econometric Reviews*, **26**, 329-363.

[22] Engle, R.F. and C.W.J. Granger (1987). "Cointegration and Error-Correction: Representation, Estimation, and Testing". *Econometrica*, **55**, 251-276.

[23] Forni M., M. Hallin, M. Lippi and L. Reichlin (2005). "The Generalized dynamic factor model: One-sided estimation and forecasting". *Journal of the American Statistical Association*, **100**, 830-840.

[24] Garcia-Ferrer, A., R.A. Highfield, F. Palm and A. Zellner (1987). "Macroeconomic forecasting using pooled international data". *Journal of Business and Economic Statistics*, **5**, 53-67.

[25] Geweke J. and C. Whiteman(2006). "Bayesian Forecasting". *In Handbook of Economic Forecasting, G. Elliot , C.W.J. Granger and A. Timmermann*, North-Holland: Amsterdam; 3-80.

[26] Giannone, D., M. Lenza and G. Primiceri (2010). "Prior Selection for Vector Autoregressions". *mimeo*

[27] Giannone, D., L. Reichlin and D. Small (2008). "Nowcasting: The Real-Time Informational Content of Macroeconomic Data Releases". *Journal of Monetary Economics*, **55**, 665-676.

[28] Giannone, D., L. Reichlin and S. Simonelli (2009). "Nowcasting Euro Area Economic Activity in Real Time: The Role of Confidence Indicators". *National Institute Economic Review*, **210**, 90-97.

[29] Hamilton, J.D.(1994). "Time Series Analysis". Princeton University.

[30] Litterman, R. B. (1980). "A Bayesian Procedure for forecasting with vector autoregressions". *Working Paper, MIT*

[31] Litterman, R. B. (1984). "Forecasting and Policy Analysis with Bayesian Vector Autoregression Models". *Federal Reserve Bank of Minneapolis Quarterly Review*, **8**, 3041.

[32] Litterman, R. B. (1986). "Forecasting with bayesian vector autoregressions -5 years of experience". *Journal of Business and Economic Statistics*,**4**, 25-38.

[33] Lubik, T.A. and F. Schorfheide (2005). "A Bayesian Look at New open Economy Macroeconomics". *NBER Macroeconomics Annual*, 313-386.

[34] Robbins, H. (1956). "An Empirical Bayes Approach to Statistics". *Preceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability: Contributions to the Theory of Statistics*,**1**, 157-163.

[35] Raynauld, J. and J.G. Simonato (1993). "Seasonal BVAR models". *Journal of Econometrics*, **55**, 203-229.

[36] Sims C.A., J.HJ. Stock and M.W. Watson (1990). "Inference in Linear Time Series Models with Some Unit Roots ". *Econometrica*, **58**, 113-144.

[37] Sims C.A. (1993). "A Nine Variable Probabilistic Macroeconomic Forecasting Model". In *J.H. Stock and M.W. Watson (eds.), Business Cycle Indicators and Forecasting*, University of Chicago Press for the NBER, 179-204.

[38] Sims C.A., T. Zha (1998). "Bayesian Methods for Dynamic Multivariate Models". *International Economic Review*, **39**, 949-68.

[39] Stock J.H., and M.W. Watson (2002). "Forecasting using principal components from a large number of predictors". *Journal of the American Statistical Association*, **97**, 147-162.

[40] Stock J.H., and M.W. Watson (1999). "A comparison of Linear and Nonlinear Univariate Models for Forecasting Macroeconomic Time Series". In *Cointegration, Causality, and Foreasting A Festschrift in Honour of Clive W.J. Granger, ed. by R.F. Engle, and H. White*, pp. 1-44. Oxford University Press, Oxford

[41] Stock J.H., and M.W. Watson (2004). "Combination forecasts of output growth in a seven-country data set". *Journal Forecasting*, **23**, 405-430.

[42] Theil, H., A. S. Goldberger (1961). "On pure and mixed estimation in econometrics". *International Economic Review*, **2**, 65-78.

[43] Timmermann, A.(2006). "Forecast Combinations". *In Handbook of Economic Forecasting, G. Elliot , C.W.J. Granger and A. Timmermann*, North-Holland: Amsterdam; 135-196.

A VAR Priors for our Dynamic Regression

A.1 VAR Models

Vector autoregressive models are flexible enough to capture the dynamic correlation patterns between GDP and all business cycle indicators. For the sake of simplicity, consider a bivariate VAR with $p = 2$. Let the first variable can be the level of GDP (Y_t) and the second variable employment (A_t):

$$\begin{bmatrix} \mathrm{Y}_t \\ \mathrm{A}_t \end{bmatrix} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} Y_{q-1} \\ A_{q-1} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} Y_{q-2} \\ A_{q-2} \end{bmatrix} + \begin{bmatrix} \upsilon_t \\ \chi_t \end{bmatrix}$$

The notation can be further simplified to

$$y_t^{'} = x_t^{'}\Theta + \epsilon_t^{'}, \quad \epsilon_t \sim N(0, \Sigma), t = 1, \ldots, T \tag{2}$$

with

$$y_q = \begin{bmatrix} Y_t \\ A_t \end{bmatrix}, x_t = \begin{bmatrix} Y_{t-1} \\ A_{t-1} \\ Y_{t-2} \\ A_{t-2} \\ \mathbf{1} \end{bmatrix}, \epsilon_q = \begin{bmatrix} \upsilon_t \\ \chi_t \end{bmatrix}, \Theta = \begin{bmatrix} \beta_{11} & \beta_{21} \\ \beta_{12} & \beta_{22} \\ \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \\ \alpha_1 & \alpha_2 \end{bmatrix}$$

In matrix notation,

$$\underbrace{Y}_{T\times 2} = \underbrace{X}_{T\times 5}\underbrace{\Theta}_{5\times 2} + \underbrace{E}_{T\times 2} \tag{3}$$

Note that for a larger number of variables and a larger p the number of parameters increases dramatically, which generally guarantees a good in-sample fit. However, the *inefficient* estimation of a large number of parameters will give as a result poor out-of-sample projections.

A.2 Prior Design

Linear combinations of parameters defined within and accross VAR equations will be shrunk in accordance with statistical knowledge that is common to most macroeconomic data. For example, we will impose the prior that there are unit roots in the individual series, letting the data define whether those unit roots are driven by few stochastic trends or by independent trends. The common practice of taking (exact) differences to stationarize the data implies that the VAR representation would be misspecified in the presence of co-integration[15].

Here, our prior believes enter the system through dummies or artificial observations of Y and X. This is often interpreted as *mixed estimation* since Theil and Goldberger (1961). Thus, the dummy observations are mixed with the actual sample according to the following simple equation:

$$\hat{\Theta} = (X^{'}X + X^{*'}X^{*})^{-1}(X^{'}Y + X^{*'}Y^{*}) \tag{4}$$

Nevertheless, a fully bayesian perspective is often considered in the literature. Such an interpretation implies that the prior distribution of the VAR parameters combines the likelihood function for the dummy observations with an improper prior $p(\Theta, \Sigma) \propto |\Sigma|^{-(N+1)/2}$, where N is the dimension of the VAR. Doan, Litterman and Sims (1984) or Sims and Zha (1998) provide a detailed exposition.

The first two dummies described below instrumentalize the so-called Minnesota prior (see Litterman, 1980), while the next two types of dummies contribute to imposing independent beliefs about the presence of unit roots and co-integration (see Sims and Zha, 1998). Those priors are parameterized here though τ, λ, μ, and d, following **Lubik and Schorfheide (2005)**. Later, we will explain how to choose values for those parameters. For the time being, we

[15] Since Engle and Granger (1987), it has been common to estimate VARs through the introduction of error correction mechanisms. However, estimation of VARs in levels is also possible independently of the order of integration of the series and the number of co-integration relationships. See for example Sims, Stock and Watson (1990)

assume that those values are given:

1. **Dummies for the coefficients associated to the first lag**

   Consider equation (2). For our simple bivariate VAR with two lags, the dummy observations take the following form:

$$\underbrace{\begin{bmatrix} \tau s_1 & 0 \\ 0 & \tau s_2 \end{bmatrix}}_{dummy\ \text{“}observations\text{''}\ Y^*} = \underbrace{\begin{bmatrix} \tau s_1 & 0 & 0 & 0 & 0 \\ 0 & \tau s_2 & 0 & 0 & 0 \end{bmatrix}}_{dummy\ \text{“}observations\text{''}\ X^*} \begin{bmatrix} \beta_{11} & \beta_{21} \\ \beta_{12} & \beta_{22} \\ \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \\ \alpha_1 & \alpha_2 \end{bmatrix} + \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

   The parameter τ is the tightness of the prior, and two terms, s_1 and s_2, capture the variance of each time series. These two dummies introduce prior knowledge into the coefficients associated with the first lag. While the “own” autoregressive coefficients are shrunk towards 1, the prior for the remaining coefficients is centered around 0. One can understand this idea by noticing the the above system of “beliefs” implies that:

$$\begin{aligned} \tau s_1 &= \tau s_1 \beta_{11} + e_{11} \Rightarrow \beta_{11} = 1 + \frac{e_{11}}{\tau s_1} \\ 0 &= \tau s_1 \beta_{21} + e_{12} \Rightarrow \beta_{21} = 0 + \frac{e_{12}}{\tau s_1} \\ 0 &= \tau s_2 \beta_{12} + e_{21} \Rightarrow \beta_{12} = 0 + \frac{e_{21}}{\tau s_2} \\ \tau s_2 &= \tau s_2 \beta_{22} + e_{12} \Rightarrow \beta_{22} = 1 + \frac{e_{22}}{\tau s_2} \end{aligned}$$

   Although the precise effect of these dummies is given by their likelihood function, the equations above suggest a heuristic explanation of the role of τ. Under the normality assumption on the error terms, τ determines the precision of the prior on the four coefficients associated to the first lag:

$$\beta_{11} \sim N\left(1, \frac{1}{\tau}\frac{\sigma_{11}}{s_1}\right)$$
$$\beta_{21} \sim N\left(0, \frac{1}{\tau}\frac{\sigma_{21}}{s_1}\right)$$
$$\beta_{12} \sim N\left(0, \frac{1}{\tau}\frac{\sigma_{21}}{s_2}\right)$$
$$\beta_{22} \sim N\left(1, \frac{1}{\tau}\frac{\sigma_{22}}{s_2}\right)$$

2. **Dummies for the coefficients associated to the second lag** ($p = 2$)

$$\underbrace{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}}_{dummy\ \text{``}observations\text{''}\ Y^*} = \underbrace{\begin{bmatrix} 0 & 0 & \tau s_1 p^d & 0 & 0 \\ 0 & 0 & 0 & \tau s_2 p^d & 0 \end{bmatrix}}_{dummy\ \text{``}observations\text{''}\ X^*} \begin{bmatrix} \beta_{11} & \beta_{21} \\ \beta_{12} & \beta_{22} \\ \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \\ \alpha_1 & \alpha_2 \end{bmatrix} + \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

These dummies shrink *all* the autoregressive coefficients associated with the second (and subsequent) lag(s) towards 0. The tightness of the prior is given by τ, as in the previous case, and by p^d. Thus, the parameters associated with more distant lags are more strongly shrunk towards 0.

3. **Co-persistence** As opposed to the previous two priors, this one does not aims to impose beliefs about individual coefficients but linear combinations of them. This prior takes the form of a single observation of the VAR system:

$$\underbrace{\begin{bmatrix} \lambda\bar{y}_1 & \lambda\bar{y}_2 \end{bmatrix}}_{dummies\ \text{``}observations\text{''}\ Y^*} = \underbrace{\begin{bmatrix} \lambda\bar{y}_1 & \lambda\bar{y}_2 & \lambda\bar{y}_1 & \lambda\bar{y}_2 & \lambda \end{bmatrix}}_{dummy\ \text{``}observations\text{''}\ X^*} \begin{bmatrix} \beta_{11} & \beta_{21} \\ \beta_{12} & \beta_{22} \\ \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \\ \alpha_1 & \alpha_2 \end{bmatrix} + \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{bmatrix}$$

This prior is also called "dummy initial observation" or "one-unit-root prior". This dummy adds to the likelihood the following term, which has more weight for large values of λ (the parameter governing the tightness of this prior):

$$-\frac{1}{2}log|\Sigma| - \frac{\lambda^2}{2}\left(\underbrace{(I-B-\Gamma)\overline{y}-\alpha}_{innovation}\right)'\Sigma^{-1}\left(\underbrace{(I-B-\Gamma)\overline{y}-\alpha}_{innovation}\right)$$

where

$$\alpha=\begin{bmatrix}\alpha_1\\ \alpha_2\end{bmatrix}, \overline{y}=\begin{bmatrix}\overline{y}_1\\ \overline{y}_2\end{bmatrix}, \text{B}=\begin{bmatrix}\beta_{11} & \beta_{12}\\ \beta_{21} & \beta_{22}\end{bmatrix}, and\ \ \Gamma=\begin{bmatrix}\gamma_{11} & \gamma_{12}\\ \gamma_{21} & \gamma_{22}\end{bmatrix},$$

and $\overline{y}$ is chosen to be equal to the mean of the first observations.

The particularity of this dummy observation is that it imposes a bimodal prior distribution on the VAR coefficients. The prior favours on the one hand the area of the parameter space where $\alpha = 0$ and the system contains at least one unit root $|I - B - \Gamma| = 0$. Second, the prior density also concentrates on regions where $\alpha \neq 0$ and y_t is stationary. This attributes some probability to the possibility that the initial observations (or its average $\overline{y}$) are close to the unconditional mean of the model. The combination of this prior with the next one, which favours the presence of stochastic trends, may help to provide convenient beliefs for the estimation of our VARs in levels.

4. **Own-persistence**

$$\underbrace{\begin{bmatrix}\mu\overline{y}_1 & 0\\ 0 & \mu\overline{y}_2\end{bmatrix}}_{dummies\ ``observations''\ Y^*} = \underbrace{\begin{bmatrix}\mu\overline{y}_1 & 0 & \mu\overline{y}_1 & 0 & 0\\ 0 & \mu\overline{y}_2 & 0 & \mu\overline{y}_2 & 0\end{bmatrix}}_{dummy\ ``observations''\ X^*}\begin{bmatrix}\beta_{11} & \beta_{21}\\ \beta_{12} & \beta_{22}\\ \gamma_{11} & \gamma_{21}\\ \gamma_{12} & \gamma_{22}\\ \alpha_1 & \alpha_2\end{bmatrix} + \begin{bmatrix}e_{11} & e_{12}\\ e_{21} & e_{22}\end{bmatrix}$$

This type of dummies are widely used in the literature. They contribute to the incorporation of the belief that there is no co-integration in the system. The precision of this prior is given by μ, . However, this does not amount to ruling out the presence of of co-movements in our data, since it only restricts linear combinations of the coefficients. This approach is often known as "inexact differencing".

The following error correction representation of our illustrative bivariate VAR(2) helps us to understand the implications of this prior.

$$\Delta y_t = \alpha - \underbrace{(I_2 - B - \Gamma)}_{co-integration} y_{t-1} - B\Delta y_{t-1}$$

By shrinking $(I_2 - B - \Gamma)$ towards zero, the prior mitigates the cointegration relationships. Nevertheless, this does not necessarily mean that the variables in y_t do not co-move in *long-run* frequencies, since the posterior distribution will also be affected by the likelihood function of the data. Moreover, since the coefficients of B and Γ are not individually shrunk to zero, but the prior is over *sums of coefficients*, a strong shrinkage towards zero would not be able to cancel the ability of the parameters to capture *short run* co-movements.

5. **Prior on the covariance matrix** The dummies for the covariance matrix of the error terms, one for each equation of the VAR, take the following form:

$$\underbrace{\begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix}}_{dummy\ ``observations'' \ Y^*} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{dummy\ ``observations'' \ X^*} \begin{bmatrix} \beta_{11} & \beta_{21} \\ \beta_{12} & \beta_{22} \\ \gamma_{11} & \gamma_{21} \\ \gamma_{12} & \gamma_{22} \\ \alpha_1 & \alpha_2 \end{bmatrix} + \begin{bmatrix} e_{11} & e_{12} \\ e_{21} & e_{12} \end{bmatrix}$$

We fix σ_i equal to the standard deviation of the first observations of each variable i. On the other hand, $\overline{y}_i$ is equal to the sample mean of the initial observations.

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS[1]

0922 GALO NUÑO: Technology, convergence and business cycles.
0923 FRANCISCO DE CASTRO AND JOSÉ LUIS FERNÁNDEZ: The relationship between public and private saving in Spain: does Ricardian equivalence hold?
0924 GONZALO FERNÁNDEZ-DE-CÓRDOBA, JAVIER J. PÉREZ AND JOSÉ L. TORRES: Public and private sector wages interactions in a general equilibrium model.
0925 ÁNGEL ESTRADA AND JOSÉ MANUEL MONTERO: R&D investment and endogenous growth: a SVAR approach.
0926 JUANA ALEDO, FERNANDO GARCÍA-MARTÍNEZ AND JUAN M. MARÍN DIAZARAQUE: Firm-specific factors influencing the selection of accounting options provided by the IFRS: Empirical evidence from Spanish market.
0927 JAVIER ANDRÉS, SAMUEL HURTADO, EVA ORTEGA AND CARLOS THOMAS: Spain in the euro: a general equilibrium analysis.
0928 MAX GILLMAN AND ANTON NAKOV: Monetary effects on nominal oil prices.
0929 JAVIER MENCÍA AND ENRIQUE SENTANA: Distributional tests in multivariate dynamic models with Normal and Student *t* innovations.
0930 JOAN PAREDES, PABLO BURRIEL, FRANCISCO DE CASTRO, DANIEL GARROTE, ESTHER GORDO AND JAVIER J. PÉREZ: Fiscal policy shocks in the euro area and the US: an empirical assessment.
0931 TERESA LEAL, DIEGO J. PEDREGAL AND JAVIER J. PÉREZ: Short-term monitoring of the Spanish Government balance with mixed-frequencies models.
0932 ANTON NAKOV AND GALO NUÑO: *Oil*gopoly: a general equilibrium model of the oil-macroeconomy nexus.
0933 TERESA LEAL AND JAVIER J. PÉREZ: Análisis de las desviaciones presupuestarias aplicado al caso del presupuesto del Estado.
0934 JAVIER J. PÉREZ AND A. JESÚS SÁNCHEZ: Is there a signalling role for public wages? Evidence for the euro area based on macro data.
0935 JOAN PAREDES, DIEGO J. PEDREGAL AND JAVIER J. PÉREZ: A quarterly fiscal database for the euro area based on intra-annual fiscal information.
1001 JAVIER ANDRÉS, ÓSCAR ARCE AND CARLOS THOMAS: Banking competition, collateral constraints and optimal monetary policy.
1002 CRISTINA BARCELÓ AND ERNESTO VILLANUEVA: The response of household wealth to the risk of losing the job: evidence from differences in firing costs.
1003 ALEXANDER KARAIVANOV, SONIA RUANO, JESÚS SAURINA AND ROBERT TOWNSEND: No bank, one bank, several banks: does it matter for investment?
1004 GABRIEL PEREZ-QUIROS AND HUGO RODRÍGUEZ MENDIZÁBAL: Asymmetric standing facilities: an unexploited monetary policy tool.
1005 GABRIEL JIMÉNEZ, JOSE A. LOPEZ AND JESÚS SAURINA: How does competition impact bank risk-taking?
1006 GIUSEPPE BERTOLA, AURELIJUS DABUSINSKAS, MARCO HOEBERICHTS, MARIO IZQUIERDO, CLAUDIA KWAPIL, JEREMI MONTORNÈS AND DANIEL RADOWSKI: Price, wage and employment response to shocks: evidence from the WDN Survey.
1007 JAVIER MENCÍA: Testing non-linear dependence in the Hedge Fund industry.
1008 ALFREDO MARTÍN-OLIVER: From proximity to distant banking: Spanish banks in the EMU.
1009 GALO NUÑO: Optimal research and development expenditure: a general equilibrium approach.
1010 LUIS J. ÁLVAREZ AND PABLO BURRIEL: Is a Calvo price setting model consistent with micro price data?
1011 JENS HAGENDORFF, IGNACIO HERNANDO, MARÍA J. NIETO AND LARRY D. WALL: What do premiums paid for bank M&As reflect? The case of the European Union.
1012 DAVID DE ANTONIO LIEDO: General equilibrium restrictions for dynamic factor models.
1013 JAMES COSTAIN, JUAN F. JIMENO AND CARLOS THOMAS: Employment fluctuations in a dual labor market.
1014 LUIS M. VICEIRA AND RICARDO GIMENO: The euro as a reserve currency for global investors.
1015 PALOMA LÓPEZ-GARCÍA AND JOSÉ MANUEL MONTERO: Understanding the Spanish business innovation gap: The role of spillovers and firms' absorptive capacity.

1. Previously published Working Papers are listed in the Banco de España publications catalogue.

1016 AITOR LACUESTA AND SERGIO PUENTE: El efecto del ciclo económico en las entradas y salidas de inmigrantes en España.

1017 REBEKKA CHRISTOPOULOU, JUAN F. JIMENO AND ANA LAMO: Changes in the wage structure in EU countries.

1018 THOMAS BREUER, MARTIN JANDAČKA, JAVIER MENCÍA AND MARTIN SUMMER: A systematic approach to multi-period stress testing of portfolio credit risk.

1019 LUIS J. ÁLVAREZ AND PABLO BURRIEL: Micro-based estimates of heterogeneous pricing rules: The United States vs. the euro area.

1020 ALFREDO MARTÍN-OLIVER AND VICENTE SALAS-FUMÁS: I.T. investment and intangibles: Evidence from banks.

1021 LUISA LAMBERTINI, CATERINA MENDICINO AND MARIA TERESA PUNZI: Expectations-driven cycles in the housing market.

1022 JULIÁN MESSINA, PHILIP DU CAJU, CLÁUDIA FILIPA DUARTE, NIELS LYNGGÅRD HANSEN AND MARIO IZQUIERDO: The incidence of nominal and real wage rigidity: an individual-based sectoral approach.

1023 ALESSIO MORO: Development, growth and volatility.

1024 LUIS J. ÁLVAREZ AND ALBERTO CABRERO: Does housing really lead the business cycle?

1025 JUAN S. MORA-SANGUINETTI: Is judicial inefficiency increasing the house property market weight in Spain? Evidence at the local level.

1026 MAXIMO CAMACHO, GABRIEL PEREZ-QUIROS AND PILAR PONCELA: Green shoots in the Euro area. A real time measure.

1027 AITOR ERCE AND JAVIER DÍAZ-CASSOU: Creditor discrimination during sovereign debt restructurings.

1028 RAFAEL REPULLO, JESÚS SAURINA AND CARLOS TRUCHARTE: Mitigating the pro-cyclicality of Basel II.

1029 ISABEL ARGIMÓN AND JENIFER RUIZ: The effects of national discretions on banks.

1030 GABRIEL JIMÉNEZ, STEVEN ONGENA, JOSÉ-LUIS PEYDRÓ AND JESÚS SAURINA: Credit supply: identifying balance-sheet channels with loan applications and granted loans.

1031 ENRIQUE MORAL-BENITO: Determinants of economic growth: A Bayesian panel data approach.

1032 GABE J. DE BONDT, TUOMAS A. PELTONEN AND DANIEL SANTABÁRBARA: Booms and busts in China's stock market: Estimates based on fundamentals.

1033 CARMEN MARTÍNEZ-CARRASCAL AND JULIAN VON LANDESBERGER: Explaining the demand for money by non-financial corporations in the euro area: A macro and a micro view.

1034 CARMEN MARTÍNEZ-CARRASCAL: Cash holdings, firm size and access to external finance. Evidence for the euro area.

1035 CÉSAR ALONSO-BORREGO: Firm behavior, market deregulation and productivity in Spain.

1036 OLYMPIA BOVER: Housing purchases and the dynamics of housing wealth.

1037 DAVID DE ANTONIO LIEDO AND ELENA FERNÁNDEZ MUÑOZ: Nowcasting Spanish GDP growth in real time: “One and a half months earlier”.

BANCO DE **ESPAÑA**
Eurosistema

Unidad de Publicaciones
Alcalá, 522; 28027 Madrid
Telephone +34 91 338 6363. Fax +34 91 338 6488
E-mail: publicaciones@bde.es
www.bde.es