Hyper-Parameterised Dynamic Regressions for Nowcasting Spanish GDP Growth in Real Time

David de Antonio Liedo*

Elena Fernandez Muñoz†

12th January, 2016

Abstract

Monitoring the ongoing developments in the Spanish business cycle requires the use of accurate forecasting tools that can handle a broad number of economic indicators. This paper analyses the short-term forecasting performance of hyper-parameterised dynamic regression models based on a large number of variables in levels, and compares it with state-of-the-art methods for nowcasting. Our method requires the estimation of many parameters if we wish to construct projections conditional on a large information set. The so-called "curse of dimensionality" is overcome here with prior information originating in the Bayesian VAR literature. The real-time forecast simulation conducted over the most severe phase of the Great Recession shows that our method yields reliable real GDP growth predictions almost one and a half months before the official figures are published. The usefulness of our approach is confirmed in an genuine out-of-sample evaluation that considers the period of the European sovereign debt crisis and the subsequent recovery.

Keywords: Bayesian shrinkage, co-movements, mixed estimation, prior elicitation, dynamic factor models, nowcasting plugin *J*Demetra+.

JEL codes: C32, C53, E37

*E-mail:david.deantonioliedo@nbb.be. DISCLAIMER: The views expressed in this paper are the authors' and not those of Banco de España or the National Bank of Belgium.

†I.E.S. Brianda de Mendoza

Contents

1 Introduction

After ten years of stable growth slightly below 4% per year in Spain, the 2008 recession has provided an excellent opportunity to "*stress-test*" our non-judgemental nowcasting models in real time. The term nowcasting made popular by Giannone *et al.* (2008) and Evans (2005) refers to the prediction of the most recent past, the present, and the nearest future[1].

In this paper, we evaluate the accuracy of out-of-sample nowcasts of Spanish GDP growth that have been constructed by regressing the log level of real GDP on its own lags and current and past values of a large and heterogeneous number of indicators. Since those dynamic regression models are hyper-parameterised in the sense that there are too many parameters to estimate, this may lead to very volatile forecasts. This paper proposes a method to improve the accuracy of these large dynamic regression models. The method exploits prior information regarding the statistical properties of the series to enhance the process of estimation. Thus, our early estimates of growth are obtained combining Bayesian priors with information *sub*sets available to the forecasters one and a half months before the official GDP figure is released by the statistical agency[2]. The forecasting evaluation has been executed with real-time data, which simulates the actual environment of professional forecasters, following a practice that has gradually become standard since the work by Croushore and Stark (2001).

Some of the existing tools available for nowcasting Spanish GDP growth take into account the presence of strong co-movements in macroeconomic data by incorporating restrictions inspired by the literature on dynamic factor models. Thus, Camacho and Domenech (2012), Camacho and Pérez-Quirós (2011) or Cuevas and Quillis (2010) propose factor models in order to have a parsimonious

[1]see Banbura *et al.* (2011) and Banbura *et al.* (2013) for a recent overview of the literature.

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

representation of GDP growth, which is 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 the euro area data, e.g. Angelini *et al.* (2011) or Camacho and Pérez-Quirós (2010), and in many other countries.

In contrast to the idea of having parsimonious models, we exploit large hyper-parameterised dynamic regression models to build GDP projections conditional on subsets of indicators aggregated at the quarterly frequency. Time series with missing observations are shifted forward in order to obtain a balanced end-of-sample structure. This approach is different from the mixed data-frequency sampling (MIDAS) regressions with leads proposed by Clements and Galvao (2008) because the simplicity of our aggregation scheme combined with our prior beliefs will allow us to incorporate information from a much larger set of variables. Our approach also differs from the class of dynamic factor models cited above, since we do not *impose* the presence of co-movements by reducing the available monthly information into one or a few quarterly factors. The potential multicollinearity problems arising from the high degree of synchronisation 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.* (2010) 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" method based on hyper-parameterised Bayesian dynamic regression models. The models we consider allow us to obtain GDP nowcasts conditional on the first p lags and the current and past values of a set of N indicator variables aggregated to the quarterly frequency. This can be considered to be a simplified version of the mixed-frequency VAR for monthly data proposed by Schorfheide and Song (2015). In both cases, 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 curse of dimensionality is tackled by using Minnesota-type priors on the coefficients, which has been standard practice since Litterman (1980, 1984,1986). The main problem of this approach is that there is no literature regarding the choice of the parameters that determine the strength of the prior believes with Spanish data.

Our real-time forecasting evaluation 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 the values that yield an optimal out-of-sample performance over a pre-sample or *training sample*. The small block of parameters with the privilege of conditioning the others are often called hyper-parameters, which is the second reason why we argue it is natural to call our models hyper-parameterised regressions. Other related approaches have been proposed in the literature. Banbura *et al.* (2010), for example, determine the strength of the prior beliefs by restricting the in-sample fit in the context of large Bayesian VARs in order to avoid over-fitting. More recently, Schorfheide and Song (2015) propose to maximize the marginal likelihood of the data, which cannot be obtained analytically in the context of missing observations. Finally, a more sophisticated approach is proposed by Giannone *et al.* (2015), who parameterize the strength of the prior beliefs and estimate them simultaneously with the rest of the parameters.

A second key element of our approach is that we take into account model uncertainty with the aim of having a deeper understanding of the use of hyper-parameterised regressions. 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 smallest forecasting errors throughout the training sample or, alternatively, equal weights for all models. Interestingly, we will also assess how the forecasts deteriorate when given indicators are excluded from the analysis, which will prove to be very useful to understand the robustness of the results.

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 factor models: considerable comovements over the business cycle, and the presence of measurement errors. In addition to that, we find that larger regressions appear to have a superior forecasting performance, presumably due to a reduction in the risk of model misspecification. In the context of nowcasting, model misspecification can also be understood as the omission of timely indicators and the inclusion of redundant predictors available with long publication lags. The disadvantage of our univariate approach with respect to the use of a unique multivariate model is that it becomes difficult to determine precisely how each one of the indicators contributes to forecasting GDP (see Banbura and Runstler, 2011). This represents a serious drawback from the point of view of understanding the forecast and communicating it. As shown by Banbura *et al.* (2011) and Banbura *et al.* (2013), any model written in state-space can decompose the forecast revisions in terms of the intraquarterly publication of "news". This type of analysis is a fundamental aspect of nowcasting, but it is unfortunately not possible within the context of linear regression models, including bridge models, regressions on factors and MIDAS regressions.

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 describes the method proposed including all concepts required for a Bayesian interpretation of our estimation procedure. 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 with a special focus on the Great Recession period as well as an evaluation of alternative *ex-ante* forecasting strategies, which are compared to a survey of professional forecasters. Section 5 incorporates additional data realisations over the second recession, driven by the European sovereign debt crisis, and the subsequent recovery. Our aim is to compare the forecasts of our hyper-parameterised dynamic regressions to those resulting from state-of-the-art Dynamic Factor Models estimated with JDemetra+. The last section concludes.

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 for the second quarter was published by the statistical agency, monthly employment figures and various surveys corresponding to April, May and June were already available. Other important variables such as sales and industrial production were also available only for April and May. Finally, real exports and imports were available until April. The complete list of variables used and publication lags can be found in Table 2.

[INSERT Table 1 here]

[INSERT Table 2 here]

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

2.1 Hyper-Parameterized Regressions

The dataset 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. All our variables are available at a monthly frequency, except for GDP growth, which is a quarterly variable. The strategy followed in this paper to deal with missing observations and mixed frequencies circumvents the problem of extracting a monthly GDP signal, as implicitly done by Banbura and Modugno (2014) or by Camacho and Pérez-Quirós (2010). Time series subject to publication lags are shifted forward in order to obtain a balanced end-of-sample structure before aggregating them to the quarterly frequency. Then, quarterly GDP is expressed as a linear function of the resulting quarterly indicators. This method is simple and has the potential advantage of eliminating part of the noise in monthly information. As suggested by Armesto *et al.* (2010), this is a valid strategy to mix frequencies without the need to specify all the variables in their original frequency. The main advantage of this method is that it does not require using additional models to fill in the missing observations at the end of the sample, as typically done by practitioners and users of bridge equation models. Other aggregation schemes are possible. For example, Clements and Galvao (2008) propose to use the mixed data-frequency sampling (MIDAS) methodology of Ghysels *et al.* (2004) and Ghysels *et al.* (2006) to parameterise and estimate a direct link between the variables to be predicted and all the leading indicators available in their original frequency. We discarded this approach because it multiplies the number of parameters per predictor variable and it adds an extra layer of complexity to the estimation of our large regressions. Andreou *et al.* (2013) solve this problem by combining many small models, but such an approach would not allow us to explore the potential accuracy gains in the use of large models, which is the main goal of this paper.

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:

$$
\begin{aligned}
P\left(Y_t|\Omega\right) = \hat{\alpha} \quad & + \quad \hat{\beta}_{11} Y_{t-1} \; + \; \sum_{i=1}^{N} \hat{\beta}_{1,1+i} \; X_{i,t} \\
& + \quad \hat{\delta}_{11} Y_{t-2} \; + \; \sum_{i=1}^{N} \hat{\delta}_{1,1+i} \; X_{i,t-1} \\
& + \quad \ldots \\
& + \quad \hat{\gamma}_{11} Y_{t-p} \; + \; \sum_{i=1}^{N} \hat{\gamma}_{1,1+i} \; X_{i,t-p+1} \qquad (1)
\end{aligned}
$$

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 obtained with the mixed estimation approach of Theil and Goldberger (1961), which is described below. Essentially, sample information is mixed with dummy observation priors that reflect the presence of unit roots in the data.

The underlying regression model can be represented by stacking all indicators on the right-hand side:

$$Y_t = X_t^{'}\Theta + \upsilon_t, \qquad \upsilon_t \sim N(0, \sigma_\upsilon^2), t = 1, \ldots, T \tag{2}$$

with

$$X_t = \left[\begin{array}{c} Y_{t-1} \\ X_{1,t} \\ \vdots \\ X_{N,t} \\ \hline Y_{t-2} \\ X_{1,t-1} \\ \vdots \\ X_{N,t-1} \\ \hline \vdots \\ \hline 1 \end{array}\right], \Theta = \left[\begin{array}{c} \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1N} \\ \hline \gamma_{11} \\ \gamma_{12} \\ \vdots \\ \gamma_{1N} \\ \hline \vdots \\ \hline \alpha \end{array}\right]$$

Note that for a larger number of variables and a larger p the number of parameters required by expression (1) increases dramatically, which leads to very inefficient estimates. This problem of having too many parameters could be solved by adding additional observations, which we do not have. However, an equivalent solution is the use of the so-called dummy observations or artificial data which will be eventually interpreted as prior beliefs. Such an approach will be fully understood using a matrix notation to define the regression model for our data:

$$\underbrace{Y}_{T\times 1} = \underbrace{X}_{T\times k}\underbrace{\Theta}_{k\times 1} + \underbrace{E}_{T\times 1}, \tag{3}$$

with $k = (N+1)p + 1$. In a simple example where we only use one indicator to forecast GDP and two lags, $k = 5$.

2.2 Bayesian Estimation

The number of parameters in equation 2 defining the link between our time series of interest Y and the potentially large set of predictor variables X may be large. Unless the parameter space can be restricted, this dimensionality problem reduces the quality of the forecasts due to a decrease in the degrees of freedom. One solution would be to use variable selection procedures that result in a lower number of variables. As discussed by De Mol *et al.* (2008), data selection procedures such as the Lasso (i.e. least absolute shrinkage and selection operator) can be interpreted as the result of introducing a prior in the form of a density with large probability mass around zero and the tails for all parameters, to make sure the variables are either selected or rejected. However, under collinearity, this variable selection technique is likely to be very unstable, since distinct subsets composed by carefully chosen number of predictor variables may yield equivalent forecasts. Because the business cycle indicators are characterized by strong collinearity, these variable 'selection' techniques are unlikely to provide any added value with respect to variable 'aggregation' methods, such as the penalised Ridge regressions and regressions on principal components[3]. In contrast to De Mol *et al.* (2008), who focus on the static linear regression framework with a large number of predictor variables, our approach involves dynamic regressions.

The dynamic nature of our regression equation suggests a slightly different approach. We propose to use prior beliefs that represent statistical knowledge regarding the dynamics of macroeconomic time series data. For example, one can impose as a prior belief that there are unit roots in the individual series, letting the data define whether those unit roots are driven by only few stochastic trends.

Here, our prior beliefs enter the system through dummies or artificial observations that are added to the T rows of Y and X in expression (3). This has

[3] Hastie *et al.*(2001), Section 3.6, discuss the relationship between principal components and Bayesian regressions.

often been 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:

$$\hat{\Theta} = (X'X + X^{0\prime}X^{0})^{-1}(X'Y + X^{0\prime}Y^{0}) \tag{4}$$

A fully Bayesian perspective can be considered by constructing a prior distribution that combines the likelihood function for the dummy observations with an improper prior $p(\Theta, \sigma_{\upsilon}) \propto |\sigma|^{-2}$ for all the parameters. Doan, Litterman and Sims (1984) or Sims and Zha (1998) provide a detailed description when this idea is applied to a VAR model.

2.3 Prior Elicitation

Bauwens *et al.* (2003, Chapter 4: "The quantification of ignorance") provide a detailed overview of how the prior elicitation problem has been solved in the context of linear regression models like the one we use in this paper. We will first clarify the choice of the prior density before we explain in detail how we determine the value of the hyperparameters that govern the tightness or precision of the prior.

The form of the prior is based on the *natural conjugate* principle. Natural conjugate priors have the same functional form as the likelihood function. They lead to posterior densities with sufficient statistics that can be written as a combination of the sufficient statistics for the prior and likelihood functions. Raiffa and Schalaifer (1961) suggest that, within the exponential family, the likelihood kernel of a hypothetical previous sample could be interpreted as the natural conjugate prior.

It is well known that the likelihood function of the linear regression model defined in equation 3 is proportional to the kernel of a normal-inverted gamma−2 density represented as $NIG(\underbrace{\mu}_{\hat{\Theta}}, \underbrace{M}_{X'X}, s, \underbrace{\nu}_{T-k-2})$ in Θ and σ_{υ}^{2}, which is described below[4]:

$$L(\Theta,\sigma_v^2|Y,X) \propto (\sigma_v^2)^{(-T/2)} \exp\left(-\frac{1}{2\sigma_v^2}\left[\underbrace{s+(\Theta-\hat{\Theta})'X'X(\Theta-\hat{\Theta})}_{(Y-X\Theta)(Y-X\Theta)'}\right]\right) \tag{5}$$

where

$$\hat{\Theta} = (X'X)^{-1}X'Y \tag{6}$$

$$s = Y'(Y - X\hat{\Theta}) \tag{7}$$

Thus, the natural conjugate prior for in Θ and σ_v^2 is a $NIG(\Theta_0, M_0, s_0, \nu_0)$ density. This can be multiplied by the likelihood kernel (5) in order to obtain the kernel of the posterior density. As shown for example in Bauwens *et al.* (2013, Chapter 2), the result is the kernel of a $NIG(\Theta_*, M_*, s_*, \nu_*)$:

$$\varphi_p(\Theta,\sigma_v^2|Y,X) \propto (\sigma_v^2)^{-(\nu_*+k+2)/2/*} \exp\left(-\frac{1}{2\sigma_v^2}\left[s_* + (\Theta-\Theta_*)'(M_0+X'X)(\Theta-\Theta_*)\right]\right) \tag{8}$$

where

$$s_* = s_0 + s + \Theta_0' M_0 \Theta_0 + \hat{\Theta}' X'X\hat{\Theta} + \Theta_*' M_* \Theta_* \tag{9}$$

$$\Theta_* = (M_0 + X'X)^{-1}(M_0\Theta_0 + X'X\hat{\Theta}) \tag{10}$$

$$\nu_* = \nu_0 + T \tag{11}$$

The last expression suggests that the posterior degrees of freedom are the result of adding the prior degrees of freedom to the sample size.

Before explaining precisely how all the prior parameters, i.e. those with a zero subindex, are determined, it is worth emphasising that the resulting posterior precision matrix is equal to the sum of the prior precision M_0 and the sample precision $X'X$. In addition to that, the posterior expectation for Θ_* is a combination of the prior mean Θ_0 and the least squares estimator $\hat{\Theta}$, and the

[4]See for example Theorem 2.22 of Bauwens *et al.* (2003). Θ and σ_v^2 have a NIG density if and only if $\Theta|\sigma_v^2 \sim N(\mu, \sigma_v^2 M^{-1})$ and $\sigma_v^2 \sim IG2(s,\nu)$.

weight of Θ_0 is given by the share of M_0 in the posterior precision. It is useful to understand what would be the kernel of the posterior density when M_0 and s_0 are equal to zero, that is, when we want the prior to be non-informative:

$$\kappa_p(\Theta|\sigma_v^2) = \sigma_v^k \exp\left[-\frac{1}{2\sigma_v^2}(\Theta - \Theta_0)' M_0(\Theta - \Theta_0)\right] \overset{M_0 \to 0}{\to} \sigma_v^{-k} \tag{12}$$

$$\kappa_p(\sigma_v^2) = \sigma_v^{\nu_0+2} \exp\left(\frac{s_0}{2\sigma_v^2}\right) \overset{s_0 \to 0}{\to} \sigma_v^{\nu_0+2} \tag{13}$$

This implies that the kernel of a diffuse prior on Θ and σ_v^2 is equal to:

$$\kappa_p(\Theta, \sigma_v^2) = \sigma_v^{\nu_0+k+2} \tag{14}$$

By choosing the hyperparameter $\nu_0 = -k$, we have the so-called Jeffrey's prior:

$$\kappa_p(\Theta, \sigma_v^2) = \sigma_v^2, \tag{15}$$

which turns out to yield a proper posterior density for Θ if the number of observations is larger than the size of Θ, i.e. $T \geq k$. This choice of $\nu_0 = -k$ also guarantees that the posterior variance for Θ increases with k:

$$Var(\Theta|X, Y) = \frac{s}{T-k-2}(X'X)^{-1}$$

2.4 Prior Hyperparameters

The hyperparameters defining the exact shape of our prior distribution will be mapped to sufficient statistics from an imaginary sample (X^0, Y^0) that has been generated, independently, by the model representing our actual observables X, Y. Those artificial observations implement a natural conjugate prior, which can be interpreted as the posterior density of that sample analysed under the non-informative prior defined above (Jeffrey). This implies that the degrees of freedom corresponding to the prior distribution increase with the size of the imaginary sample $\nu_0 = T_0 - k$. The prior mean, scale and precision parameters will be given by:

$$\Theta_0 = (X^{0'}X^0)^{-1}X^{0'}Y^0 \tag{16}$$

$$s = Y^{0'}(Y^0 - X^0\Theta_0) \tag{17}$$

$$M_0 = (X^{0'}X^0)^{-1}X^{0'}Y^0 \tag{18}$$

The first two dummies described below instrumentalise 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 defined here through the hyperparameters τ, λ, μ, and d, following Lubik and Schorfheide (2005). Alternative methods to determine those hyperparameters have been proposed by Giannone *et al.* (2015), Banbura *et al.* (2010) or Schorfheide and Song (2015). We will simply select the values that yield an optimal out-of-sample performance over a pre-sample or *training sample*[5] .

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

   Consider equation 2. For our simple regression with two lags, the dummy observations to be added to the T rows of expression (3) take the following form:

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

[5] An alternative approach would be to optimise over the predictive density of the last available GDP time series at each point in time, which is multivariate Student (see Theorem 2.25 of Bauwens, 2003), but this would exclude from our search all models with improper densities. Most importantly, we thought there could be gains from the use of external information such as a time series of Flash releases instead of the last available vintage.

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 above system of "beliefs" implies that:

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

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 with the first lag:

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

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

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

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 aim to impose beliefs about individual coefficients but linear combinations

of them. This prior also originates in the VAR literature, although in our case it takes the form of a single observation for Y_t and X_t:

$$\underbrace{\begin{bmatrix} \lambda\overline{y} \end{bmatrix}}_{dummies\ \text{``}observations\text{''}\ Y^0} = \underbrace{\begin{bmatrix} \lambda\overline{y} & \lambda\overline{x} & \lambda\overline{y} & \lambda\overline{x} & \lambda \end{bmatrix}}_{dummy\ \text{``}observations\text{''}\ X^0} \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \gamma_{11} \\ \gamma_{12} \\ \alpha_1 \end{bmatrix} + \begin{bmatrix} v \end{bmatrix}$$

This prior is also called a "dummy initial observation" or a "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):

$$-\frac{1}{2}log|\sigma_v^2| - \frac{\lambda^2}{2}\left(\underbrace{(1-\beta_{11}-\gamma_{11})\overline{y} - (\beta_{12}-\gamma_{12})\overline{x} - \alpha}_{innovation}\right)^2 \sigma_v^{-2}$$

where $\overline{y}$ and $\overline{x}$ are the average of the fist observations in the sample.

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

The particularity of this dummy observation is that it allows for a prior distribution with a mode at the point in the parameter space where at least our variable of interest has a unit root, i.e. $(1-\beta_{11}-\gamma_{11}) = 0$ and $\alpha = (\beta_{12}-\gamma_{12})\overline{x}$. But it could also be that $\alpha = (1-\beta_{11}-\gamma_{11})\overline{y}-(\beta_{12}-\gamma_{12})\overline{x}$. The last expression implies that the weight of the initial observations (or their average $\overline{y}$) becomes very important at determining the value of the parameters. This well known bias towards stationarity can also be modified by combining this prior with the next one, which favours the presence of stochastic trends. This combination may provide convenient beliefs for the estimation of our linear regression models in levels.

4. **Own persistence**

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

This type of dummy has been widely used, incorporating 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 co-movements in our data, since it only restricts linear combinations of the coefficients. This approach is often known as "inexact differencing". It can be easily shown, by writing down the equations corresponding to the two dummy observations, that $(1 - \beta_{11} - \gamma_{11})$ converges to 0 when μ increases. At the same time, $\beta_{12} + \gamma_{12}$ converges to zero, which implies that the co-integration relationships among our variables are mitigated.

The use of this prior does not necessarily mean that the variables 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 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:

$$\underbrace{\begin{bmatrix} s_1 \end{bmatrix}}_{dummy\ ``observations" \ Y^0} = \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \end{bmatrix}}_{dummy\ ``observations" \ X^0} \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \gamma_{11} \\ \gamma_{12} \\ \alpha_1 \end{bmatrix} + \begin{bmatrix} e_{11} \end{bmatrix}$$

Hyperparameters: We use the first observations of our sample to define the s_i, which is the prior standard deviation of each one of the variables. The choice of $\overline{y}_i$ will be given by the sample mean of the initial observations. The remaining hyperparameters τ, λ, μ, d can be chosen to minimise the forecast errors over a training sample.

3 Design of the Forecasting Exercise

This section describes the practical use of hyper-parameterised dynamic regressions for nowcasting in a real-time context. 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. 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 order to convince the reader of the potential of our proposed methodology, we reproduce simple*ex-ante* strategies for nowcasting in real time during the most severe phase of the recession in a context where thousands of models were available. Our method will be further tested over the European sovereign debt crisis period, which the first draft of this paper pre-dates. So, section 5 of this article will contain an additional analysis where the nowcasts from the largest possible Bayesian dynamic regression model are compared to those of state-of-the-art dynamic factor models.

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

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). To the best of our knowledge, there is no real-time database with GDP figures earlier than 1995.

The real-time nature of the forecasting practice determines the design of our evaluation exercise. The indicators described in Table 2 will be seasonally adjusted in real-time using TRAMO-SEATS[8], and introduced as predictor variables in equation 1 defined in the previous section. Note 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 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.

---

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

[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. TRAMO-SEATS is nowadays supported by *J*Demetra+ , which is a time-series software developed at the National Bank of Belgium: http://www.nbb.be/jdemetra

[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 Enrique Quillis in www.mathworks.com. Evaluating the empirical success of this alternative option is left for future research.

An Empirical Bayes Approach (EB)

Rather than using subjective beliefs, Empirical Bayes (EB) methods (Robbins, 1954) use *sample* information to elicit the priors. Here, we explore a method in this vein in order to choose the values of the hyperparameters defined in Subsection 2.4. 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 as a function of the model size are already anticipated in Figure 1.

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 major drawback of the Empirical Bayes (EB) 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 optimisation 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 straight-

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

forward to know in real time when it is the right moment to switch. Therefore, we compare the EB approach described above with the alternative of setting very Diffuse Priors (DP) 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 EB method could take years[11].

[INSERT Figure 1 here]

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. Duarte and Sussmuth (2014) propose to identify the variables with the highest correlation with Spanish GDP in order to have subset of core indicators. Nevertheless, there is no guarantee that the same indicators that have helped to produce accurate forecasts for a given sample period will continue to be helpful in the future. In this paper, 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 that contain variables that are routinely monitored by analysts of the Spanish business cycle:

[11] On average, optimising the hyperparameters to maximise 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.

- $\mathbf{\Omega_1}$: The first information set contains the 11 key variables shaded in Table 2. Those indicators provide timely information about the GDP components and the aggregate business cycle behaviour of the economy and they turn out to coincide with 8 of the variables considered in the *CF Index of Economic Activity* published by the Spanish Business Cycle Dating Committee in the website of the Spanish Economic Association. This information set includes 8 of the variables selected by Camacho and Perez-Quiros (2011): 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 one indicator of 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 transport, building permits, hotel stays, construction employment, industry PMI, the consumer confidence indicator, total sales and the 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 3 summarises the basic *ex-ante* forecasting strategies that we evaluate. With information set Ω_1, 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.

[INSERT Table 3 here]

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). As specified in the first row of Table 3, we will exploit a total of $1,023$ different models that can be constructed with Ω_1 for an assessment of how the EB and DP strategies perform in forecasting. All projections are obtained with the information available approximately one and a half months before the statistical agency publishes the national accounts.

[Insert Table 4 here]

A simple analysis of the root mean squared errors in Table 4 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 (fan chart) 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 can only be justified *ex post.*

[12]This result holds regardless of whether the forecast error is computed on the basis of the "preliminary" or the "final" GDP release. The RMSE results corresponding to the latter are included in parenthesis

[Insert Figures 2 and 3]

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 visually significant in the light of Figure 7, which considers the whole set of models. This figure compares the distribution of thousands of time series of forecast errors obtained with the EB and DP approach for the period 2008Q4-2010Q2 with the two information sets, Ω_1 and Ω_2. The two distributions are represented with a thin continuous line and a discontinuous line, respectively. They are both practically undistinguishable, specially when the errors are based on the final release (bottom panel).

[Insert Figures 4 and 5 and 7]

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 assess the performance of an alternative strategy for achieving forecasting accuracy gains. Rather than modifying our priors, we enlarge the number of predictor variables in the hope of improving forecast accuracy. As illustrated in the second row of Table 3, the larger information set Ω_2 allows us to aggregate forecasts coming from larger models. In particular, we evaluate the strategy of combining models incorporating a number of indicators ranging from ten to nineteen.

[Insert Figure 8]

The RMSE results are interesting when we compare the two subsamples of our recession episode. Table 4 provides us with overwhelming evidence in favour of Ω_2 for the second subsample, 2008Q4-2010Q2, which is visualized in Figure 8. Figure 7, which has been shown above, also displays the distribution of the errors obtained with hyper-parameterised models that result from information set Ω_2, the mode and mean of which are now closer to zero. Thus, the gains from increasing the size of the information set are actually more visible than those given by the refinement of the prior elicitation approach discussed above.

[Insert Figure 9]

On the other hand, the gradual slowdown registered over the 2006Q3-2008Q3 period has been predicted slightly more accurately with the reduced information set Ω^1, which outperforms the larger information set Ω_2 for 2007Q4 and 2008Q2. This can be seen in Figure 9, which offers a detailed picture of the forecasts. This graph also shows that the gradual slowdown registered over the 2006Q3-2008Q3 period has also been predicted accurately by the Consensus Forecast, which remains very conservative over the second subsample where both of our purely statistical models start to make a difference.

However, since 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. 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 5 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-hand 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 5 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 to the Choice of Indicators

Our relative RMSE forecasting accuracy measure is now displayed in Table 6 for 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 any significant deterioration of forecasting 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 6), the relative RMSE diminishes 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).

[Insert Table 5 and Table 6]

4.3 Interpretation of the Results

The results suggest that large regressions are more likely to identify the multiple factors underlying business cycle fluctuations, thereby reducing the risk of model misspecification and improving the quality of the forecasts over the Great Recession period.

In the context of nowcasting, model misspecification can be caused by the introduction of predictor variables with long publication lags, which can be considered to be redundant, to the detriment of more timely indicators, which are key for the early identification of turning points. Indeed, the composition effect of a large proportion of projections based on lagged information, i.e. indicators with long publication delays such as industrial production, will play a role by downweighting timely information such as the confidence and PMI indicators.

This hypothesis is supported by Figure 10, which displays the nowcasts obtained by combining parsimonious bidimensional and tridimensional projections based on the large information set Ω_2. Interestingly, the average forecast coming from those parsimonious projections on subsets of the large information set Ω_2 are highly correlated with those obtained with the small set Ω_1. This suggests that the gains of using a wider information set come from the ability to use larger models. This is not in contradiction with the fact that small-sized dynamic factor models seem to perform as well as large dynamic factor models, since the state of the art makes it possible to automatically weight the indicators depending on their timeliness and quality. The comparison of our hyper-parameterised dynamic regression with several dynamic factor models will be the subject of next section.

[Insert Figure 10]

5 Reality Check: Nowcasting 18 New Quarters of Data

According to the historical series currently available, Spanish GDP contracted by 3.6% in 2009. None of the state-of-the-art models available at that moment such as FASE, Spain-Sting or MICA-BBVA have been tested over the whole Great Recession[13]. Our idea of using hyper-parameterised dynamic regressions models was developed in 2009, with the benefit of hindsight. Whether there is any merit inherent in the proposed methodology or whether our results were driven by selecting the right indicators with the benefit of hindsight is something that could be further tested in the spirit of White (2000). However, data selection is unlikely to play a major role here. As opposed to other models such as MICA-BBVA model, for example, which introduces financial variables to improve the forecast over part of the Great Recession period that was actually triggered by a failure of the financial system, our choice of indicators follows the literature. Still, the reader may argue that the surprising performance of our hyper-parameterised regressions over the Great Recession is due to either pure luck or a data mining effort, so the real question is whether the method will continue to work during forthcoming recession episodes of unpredictable nature, without including new variables.

Five years have passed since the working paper version of our methodology was published back in 2010. The additional data, which includes the initial recovery and a second recession motivated by the European sovereign debt crisis, will allow us to conduct a genuine out-of-sample evaluation exercise. In order to sim-

[13]The evaluation period of Spain-Sting ends in 2008Q4. The last forecasts documented in the paper describing the MICA-BBVA model correspond to 2009Q1. For that period, the I.N.E. published a Flash estimate close to -2%, below the -1% figure reported by the authors of the paper. Regarding the model FASE, the only real-time results provided in the reference paper correspond to the 2009Q3-2010Q2 period, which neglects the most severe phase of the recession.

plify our reality check so that our hyper-parameterised dynamic regressions can be understood in connection with the existing methods, we will assess the results based on the largest possible regression model, which is the one that contains all indicators of the extended information set Ω_2. Therefore, we will drop from our updated analysis the idea of combining forecasts to focus on the usefulness of our approach, which requires the specification of a large number of parameters. We will compare its forecasting performance relative to dynamic factor models, which can also exploit a large number of indicators, but using a very reduced number of parameters. In particular, our results will be compared to those of a very parsimonious state-of-the-art method that automatically weights the indicators depending on their quality and timeliness: mixed-frequency stationary dynamic factor models very similar to Spain-Sting, following the methodology of Banbura and Modugno (2014) or Camacho and Pérez-Quirós (2010). The results are extremely informative, since they confirm our premise that large hyper-parameterised dynamic regressions in levels can help to improve the forecasts of state-of-the-art dynamic factor models.

5.1 Comparison of Alternative Nowcasting Methods

A comparison of several methods currently in use for nowcasting Spanish GDP growth will clarify the added value of our approach (see Table 7).

- **Hyper-parameterised dynamic regression models based on Ω_2.** Table 7 describes all the options for nowcasting Spanish GDP growth that have been discussed in this paper. The first block contains the hyper-parameterised dynamic regression models with all variables in log levels, which can be estimated either using Diffuse Priors (DP) or using the Empirical Bayes (EB) method proposed here. The last option can be executed by choosing the vector of prior tightness parameters that results in the lowest root-mean-squared errors (RMSE) over a training sample. But such an optimisation problem is solved here using either the Flash estimate (method EB_{FLASH}) or the last available figures (method EB_{LAST}) as a target to define the forecast errors. Thus, both methods are *trained* using the first and second subsamples before calculating the nowcasts for the second and third subsamples, respectively. Although it can be inferred from the sensitivity analysis of Section 4.3. that the regression based on the whole information set Ω_2 can be improved by removing a couple of variables, we will keep the whole information set, which will also be used by the competing models.

- **Mixed-Frequency Dynamic Factor Models**. The second block of nowcasting models are mixed-frequency dynamic factor models specified at a monthly frequency. They require all variables in growth rates with the exception of qualitative survey data, which is stationary at least in theory. MICA (Camacho and Domenech, 2012), Spain-Sting (Camacho and Pérez-Quirós , 2011), and FASE (Cuevas y Quillis, 2012) take into account the presence of strong co-movements in macroeconomic data by summarising all monthly indicators in terms of one pervasive factor. None of those models, which we will classify as Dynamic 1-Factor models (henceforth D(1)FM), has been reported to successfully anticipate the large declines in real activity that took place during the most severe phase of the recession, i.e. 2008Q4-2009Q2. To be fair, Spain-Sting reports a perfect nowcast for 2008Q4, and the BBVA-MICA, which was published afterwards, has produced relatively successful nowcasts also for 2009Q1. The analysis reported in the paper documenting FASE focuses on the period 2009Q3-2010Q2, thereby skipping the 2008Q4-2009Q2 period. The so-called back-testing experiment proposed in their 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.

In order to make sure the reader can finally compare our hyper-parameterised regression approach with state-of-the-art dynamic factor models, we use the *J*Demetra+ nowcasting plugin we have developed at the National Bank of Belgium. We build one model using the small information set Ω_1 and a second model exploiting the larger information set Ω_2. The two models proposed are slightly more sophisticated versions of Spain-Sting and FASE, respectively, and are estimated by combining the EM algorithm and numerical optimisation techniques for maximum likelihood in the presence of missing observations and periodic sampling. The crucial difference of our approach is that we incorporate two factors instead of one, hence the name D(2)FM. By assuming hard data loads on that second factor, the forecasts for variables such as industrial production or social security registrations improve significantly. This parameterisation downweights the excessive impact of survey data in models of this class. A second difference of our proposed factor models is that we do not specify auto-regressive dynamics in the measurement errors, which makes the forecasts less dependent on the GDP Flash releases. The simulation based on the two new models is also executed with the nowcasting plugin of *J*Demetra+, but some clarifications are needed:

- We take into account a stylized calendar for the data releases to make sure our nowcasts use the information available by the end of the reference quarter.
- The Flash GDP release and the last available time series are considered as two separate indicators. The last available GDP is assumed to be known with one year of delay, which is a convenient simplification.
- The seasonal and calendar adjustment are recursively executed with *J*Demetra+ using the TRAMO-SEATS method to extract the data

clean from seasonality, outliers and non-linearities. The recursive estimation mimics the procedure that would have been followed in real time.

5.2 RMSE: Three subsamples

The simulation exercise described considers the first two subsamples analysed in the previous sections, and a third subsample that can be used to execute a real and *not pseudo* out-of-sample validation exercise.

[**Insert here Table 8**]

5.2.1 First Subsample: 2006Q3-2008Q3

The first subsample corresponds to gradual deceleration period from 2006Q3 to 2008Q3, which is the first period with a decline in real economic activity. This subsample is used as a training sample to elicit the prior when the EB method is used, but it becomes an evaluation sample for those models that can be estimated without imposing any prior information. As shown in Table 8, both dynamic factor models beat our hyper-parameterised regression at anticipating the Flash release over this period of time. This is not surprising because such regression has been estimated with diffuse priors in the absence of a training sample, which leads to very volatile forecasts. Figure 11 illustrates that this method yields a few large forecast errors. In addition to the lack of a proper prior, our dynamic regressions face a second disadvantage. They are estimated with a panel of balanced historical data since 1995 for all series, while the dynamic factor models are based on an unbalanced panel exploiting data prior to 1995.

Not surprisingly, when those nowcasts are compared to the last available GDP growth figure, the large dynamic factor model is the only method that yields a statistically significant improvement over the random walk forecast, i.e. p-value of 0.10 suggests that the hypothesis of equal predictive accuracy can be rejected at the 90% confidence level. In order to understand those results, one can look at Figure 12 and compare the black dotted line (Flash GDP) with the shaded bars (Last GDP) in to realize that the statistical agency revised the second and third quarters of 2008 downward. Although 2008Q2 is better anticipated by dynamic factor models, they underestimate the magnitude of the large drop in economic activity that took place in the subsequent quarter.

5.2.2 Second Subsample: 2008Q4-2010Q2

The second subsample has been the main focus of our paper. It includes the quarters with the largest decline in real economic activity and ends after two consecutive quarters with positive growth rates. Such an improvement in economic activity explains why the Spanish Business Cycle Dating Committee (2015) (henceforth SBCDC) has recently decided to declare 2009Q4 as the end of the recession. We can clearly see both in Table 8 and more in detail in Figure 11 that the large volatility in the forecasts coming from the hyper-parameterised regression turns out to be useful at matching the actual volatility that was observed in the data.

This result holds regardless of whether we consider the Flash GDP or the last available estimate. The superiority of our hyper-parameterised regressions also holds when the priors are elicited on the basis of past data, but it is outperformed by the average nowcast resulting from all five different models. This result continues to support the idea that our method adds value to the already very competitive nowcasts coming from dynamic factor models.

5.2.3 Third Subsample: 2010Q3-2014Q4

The third subsample incorporates the second dip recession in the context of the European sovereign debt crisis, the peak of which has been confirmed for 2010Q4 by the SBCDC in spite of the fact that the growth rate registered in 2010Q3 is practically equal to zero. The key fact is that this critical period of time falls after the development of our hyper-parameterised regression. This provides a unique reality check to understand whether the proposed methodology remains valid on the basis of additional new evidence that was definitely unknown at the time the first version of our hyper-parameterised regressions was published as a working paper back in 2010. Table 8 suggests that both dynamic factor models and the hyper-parameterised dynamic regressions with automatically elicitated priors yield a RMSE for the Flash between 12 and 23% below the random walk benchmark on average over the subsample. Nevertheless, Figure 11 shows that the nowcasts can be very different and even contradictory in some parts of the sample, which is the reason why combining forecasts again yields a substantial improvement in forecasting accuracy of more than 40%, confirming once more the added value of our hyper-parameterised dynamic regressions.

Remarkably, the use of a diffuse prior predicted a decline in real activity of 0.34% for 2010Q4, in contrast to the 0.2% positive growth rate initially released by the I.N.E. Although the I.N.E. ended up revising heavily the whole 2014Q4-2011Q4 period , our regression models cannot exploit forthcoming revisions and hence rely on the Flash releases. This implies that when we evaluate how close the resulting nowcasts are to the last available GDP growth rates, the outcome may appear unfavourable. This explains one of the key findings in Table 8, which is the failure of the hyper-parameterised regression models to obtain statistically significant improvements over random walk forecasts for the revised figures of GDP growth. In this context, the signal extraction technology embodied in the dynamic factor models has proven to be less affected by misleading Flash estimates. In particular, the dynamic factor model based on the largest information set Ω_2 turns out to be a more accurate early estimate for the last available GDP growth rate than the Flash release itself.

[Insert Figure 12 and Figure 11]

6 Conclusion

This paper performs the most complete analysis to date of the nowcasting performance of alternative models during the most severe phase of the Great Recession. Here, we focus on the predictability of Spanish real GDP growth one and a half months before the official figures are published by the statistical agency, which corresponds to the point in time where survey data for the whole quarter is fully available. We show that hyper-parameterised dynamic regressions estimated with Minnesota type of priors turn out to yield excellent forecasts over the Great Recession period in a pseudo out-of-sample simulation exercise. Overall, our nowcasts are more accurate than the mean prediction resulting from the survey of professional forecasters published by "Consensus Economics". The good performance of our hyper-parameterised Bayesian regressions is confirmed over the period that includes the European sovereign debt recession and the subsequent recovery.

To our knowledge, our paper presents the first real-time nowcasting exercise with hyper-parameterised dynamic regression models. 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 multicollinearity problems arising from the high degree of synchronization among the predictor variables are offset by the use of priors or "inexact" restrictions originated in the VAR literature. We conclude that this method represents a valid alternative to the use of state-of-the-art dynamic factor models.

It is worth emphasizing that the Empirical Bayes methods for prior elicitation used in this paper are different from recent proposals, e.g. Giannone *et al.* (2015), Banbura *et al.* (2010) or Schorfheide and Song (2015), since we opt to increase the tightness of the prior information along its multiple dimensions in order to target our variable of interest directly rather than optimizing over the likelihood of a multivariate system or a function of it. Thus, our solution contributes to the literature that studies automatic prior elicitation strategies for nowcasting, which remains a relevant topic for further research.

7 Acknowledgements

This paper was first written while the authors were working at the "Economic Analysis and Forecasting Department" at the Banco de España. We would like to thank, without implicating, Eva Ortega, Alberto Urtasun, Samuel Hurtado and Gabriel Pérez-Quirós for comments and suggestions. We also thank the participants of the session *Forecasting Turning Points* organized by Eva Ortega at the ISF2016 (Santander) for their feedback. Finally, the authors would like to acknowledge the help of Domenico Giannone, Maximo Camacho, Andrea Silvestrini and two anonymous referees for their valuable suggestions. Section 5 of this paper has been compiled with the help of the nowcasting sofware [14] developed by the first co-author in cooperation with his current colleagues at the *R&D* Cell of the National Bank of Belgium (NBB). The software represents on of the many plugins of *J*Demetra+, which is an extensible Free and Open Source time-series software mainly developed at the NBB.

[14] Documentation and replication files: https://github.com/nbbrd/jdemetra-nowcasting/wiki

References

[1] Andreou, E., E. Ghysels, and A. Kourtellos (2013). “Should macroeconomic forecasters use daily financial data and how?”. *Journal of Business and Economic Statistics*, **31**, 240-251.

[2] Angelini, E., G. Camba-Mendez, D. Giannone, G. Runstler, and L. Reichlin (2011). “Short-Term Forecasts of Euro Area GDP Growth”. *The Econometrics Journal*, **14**, C25-C44.

[3] 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.

[4] Banbura M., and M. Modugno (2014). “Maximum likelihood estimation of factor models on data sets with arbitrary pattern of missing data”. *Journal of Applied Econometrics*, **29**, 133-160.

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

[6] Banbura M., D. Giannone and L. Reichlin (2011). “Nowcasting”. *Oxford Handbook on economic Forecasting*, in Michael P. Clements and David F. Hendy, 193-224.

[7] Banbura M., D. Giannone, M. Modugno, and L. Reichlin (2013). “Chapter 4: Now-Casting and the Real-Time Data-Flow”. *Handbook of Economic Forecasting*, in G. Elliott and A. Timmermann, eds., **2**, 195-237

[8] Banbura M., and G. Runstler (2011). “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*,**27**, 333-346

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

[10] Bauwens, L., Lubrano, M. and J.F. Richard (2003). "Bayesian Inference in Dynamic Econometric Models". *Oxford University*

[11] Camacho M., and R. Domenech (2012). "MICA-BBVA: A factor model of economic and financial indicators for short-term GDP forecasting",*SERIES: Journal of Spanish Economic Association*, **3**, 475-497.

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

[13] Camacho M., G. Perez Quiros (2011). "Spain-sting: Spain Short Term Indicator of Growth". *The Manchester School*, **79**, 594-616.

[14] Clements, M. P., and A. B. Galvao (2008). "Macroeconomic Forecasting with Mixed Frequency Data". *Journal of Business and Economic Statistics*, **26**, 546-554.

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

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

[17] Cuevas, A., E. Quillis (2012). "A Factor Analysis for the Spanish Economy (FASE)". *SERIES: Journal of Spanish Economic Association*, **3**, 311-338.

[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 R.S. Mariano (1995). "Comparing Predictive Accuracy". *Journal of Business and Economic Statistics*, **13**, 253-265.

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

[21] Duarte, P., and B. Sussmuth (2014). "Robust Implementation of a Parsimonious Dynamic Factor Model to Nowcast GDP ". *CESifo working paper no. 4574*

[22] Evans, M.D.D. (2005). "Where Are We Now? Real-Time Estimates of the Macroeconomy". *International Journal of Central Banking*, **1**, 127-175.

[23] 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.

[24] Ghysels, E., A. Sinko, and R. Valkanov (2006). "MIDAS Regressions: Further Results and New Directions". *Econometric Reviews*, **26**, 53-90.

[25] Ghysels, E., P. Santa-Clara, and R. Valkanov (2004). "The MIDAS touch: Mixed Data Sampling regression models". *mimeo*, Chapel Hill, N.C.

[26] 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.

[27] Giannone, D., M. Lenza and G. Primiceri (2015). "Prior Selection for Vector Autoregressions". *Review of Economics and Statistics*,**97**, 436-451.

[28] Hastie, T., R. Tibshirani, J. Friedman (2001). "Elements of Statistical Learning: Data Mining, Inference, and Prediction". Springer Series in Statistics.

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

[30] Litterman, R. B. (1984). “Forecasting and Policy Analysis with Bayesian Vector Autoregression Models”. *Federal Reserve Bank of Minneapolis Quarterly Review*, **8**, 30-41

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

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

[33] Raiffa, H., and R. Shaifer (1961). “Applied Statistical Decision Theory”. *Harvard Business School, Boston*, 1968 paperback edition, MIT Press, Cambridge, MA. Wiley Classics Library edition (2000)

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

[35] 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.

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

[37] Schorfheide, F., and S. Dongho (2015). “Real-Time Forecasting with a Mixed-Frequency VAR”. *Journal of Business and Economic Statistics*, **33**, 366-380.

[38] Spanish Business Cycle Dating Committee (2015), “CF Index of Economic Activity”, Spanish Economic Association.

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

[40] White, H. (2000). "A Reality Check for Data Snooping". *Econometrica*, **68**, pp. 1097-1126

Table 1: The Nowcasting Problem

Information available on 5th July 2010	NOWCASTING	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	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	current quarter	?								...
aug-10										...
sep-10										...
oct-10	next quarter	?								...
nov-10										...
dec-10										...

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**

Table 3: 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.

Table 4: Forecast accuracy with respect to the "preliminary" (and "revised") GDP

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
		(-)	(0.246)
	Diffuse Priors; DP	0.236	0.358
		(0.251)	(0.295)
Extended Information Set Ω^2 (262144 Models of size 11-20)	Diffuse Priors;DP	0.339	0.143
		(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 16% larger than the EB approach when the errors are computed on the basis of the first available GDP growth rates (16% when the errors are defined with respect to the last available vintage of GDP). An alternative option to achieve forecast accuracy is to benefit from a larger information set, Ω^2. It turns out that such a strategy provides large forecasting accuracy gains during the second subsample (2008Q4-2010Q2).

Table 5: RMSE of alternative nowcasting procedures *divided by* ***RMSE of a random walk forecast***

	Forecasting the **preliminary release**			Forecasting the **last available release**		
	full recession 2006Q3-2010Q2	start of deceleration phase 2006Q3-2008Q3	around the turning point in growth rates 2008Q4-2010Q2	full recession 2006Q3-2010Q2	start of deceleration phase 2006Q3-2008Q3	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 6: 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

Table 7: Comparison of methods to "nowcast" Spanish GDP growth in real-time

ALTERNATIVE METHODS	Number of variables	Choice of the tightness parameters	Out-of-Sample EVALUATION		
			SUBSAMPLE 1 Deceleration 2006Q3 - 2008Q3	**SUBSAMPLE 2** Recession & Recovery Part I 2008Q4-2010Q2	**SUBSAMPLE 3** Recession & Recovery Part II 2010Q3-2014Q4
Hyper reg DP (Ω_2)	20	Diffuse Prior	yes	yes	yes
Hyper reg EB_{Flash} (Ω_2)	20	Empirical Bayes (Target: Flash)	no (training sample)	yes	yes
Hyper reg EB_{Last} (Ω_2)	20	Empirical Bayeis (Target: Last)	no (training sample)	yes	yes
Mixed-Freq D(2)**FM** (Ω_2)	20	-	yes	yes	yes
Mixed-Freq D(2)**FM** (Ω_1)	10	-	yes	yes	yes
Spain-Sting D(1)**FM**	10	-	no (2008Q1-2008Q3)	no (2008Q4)	no
MICA-BBVA D(1)**FM**	12	-	yes	no (2008Q4-2009Q1)	no
FASE D(1)**FM** + transfer function (Box and Jenkins)	32	-	no	no (2009Q3-2010Q2)	no

This table provides an overview of all models cited in this paper for nowcasting Spanish real GDP growth. We also indicate the exact periods for which the forecasts of those models have been evaluated. Two remarks:

- None of the three dynamic factor models mentioned in the last three columns of this table evaluate the forcast over the whole Great Recession period, although the union of their evaluation samples does (with the exception of 2009Q2, when the negative acceleration of real activity changes its sign).

- None of the models covers the last subsample, which represents a serious challenge due to the downward revisions by the statistial agency (I.N.E.) over the first three quarters of 2011.

Table 8: RMSE for the three subsamples *divided by RMSE of Random Walk Forecasts (and Diebold-Mariano p-values)*

ALTERNATIVE METHODS	**Flash GDP**			**Last GDP**		
	SAMPLE 1 Deceleration 2006Q3 - 2008Q3	**SAMPLE 2** Recession & Recovery (I) 2008Q4-2010Q2	**SAMPLE 3** Recession & Recovery (II) 2010Q3-2014Q4	**SAMPLE 1** Deceleration 2006Q3 - 2008Q3	**SAMPLE 2** Recession & Recovery (I) 2008Q4-2010Q2	**SAMPLE 3** Recession & Recovery (II) 2010Q3-2014Q4
Hyper reg (Ω_2): DP	1.70 *(0.04)*	0.24 *(0.00)*	1.30 *(0.17)*	1.05 *(0.92)*	0.41 *(0.00)*	1.02 *(0.92)*
Hyper reg (Ω_2): EB_{Flash}	NA *(NA)*	0.77 *(0.13)*	0.79 *(0.12)*	NA (NA)	0.67 *(0.05)*	0.91 *(0.46)*
Hyper reg (Ω_2): EB_{Last}	NA (NA)	0.60 *(0.13)*	0.88 *(0.44)*	NA *(NA)*	0.58 *(0.12)*	0.97 *(0.81)*
Mixed-Freq D(2)**FM** (Ω_2)	0.77 *(0.13)*	0.91 *(0.23)*	0.84 *(0.15)*	0.66 *(0.10)*	0.79 *(0.15)*	0.58 *(0.01)*
Mixed-Freq D(2)**FM** (Ω_1)	0.66 *(0.35)*	0.83 *(0.65)*	0.77 *(0.29)*	0.88 *(0.37)*	0.73 *(0.30)*	0.73 *(0.00)*
Average all Models	1.01 *(0.97)*	0.53 *(0.01)*	0.58 *(0.00)*	0.71 *(0.45)*	0.44 *(0.01)*	0.72 *(0.01)*
I.N.E. (Revision error)	NA *(NA)*	NA *(NA)*	NA *(NA)*	0.57 *(0.21)*	0.25 *(0.02)*	0.70 *(0.03)*

The end of quarter nowcasts coming from all models are compared to either the Flash release (left-hand-side table) or the last available estimate for real GDP growth (right-hand-side table). The RMSE measure is the squared root of the average of the squared forecast errors. The last row corresponds to the revision error, which consideres the Flash release as a nowcast and compares it to the last available estimate. All RMSE measures are divided by the one resulting from the random walk benchmark. For example, the 0.70 value corresponding to the revision error over the last subsample, i.e. last column, sample 3, implies that the RMSE is 30% smaller than it would have been would the statistical agency (I.N.E.) would have released the previous Flash in place of their real-time estimates. The p-values in parentheses correspond to the Diebold-Mariano's null hypothesis∗ of equal forecasting accuracy with respect to the random walk benchmark. Note that the relative RMSE of the largest dynamic factor model over the last subsample, 0.58, is lower than that of implies that of the I.N.E. itself.

∗The Diebold-Mariano Test is defined as follows:

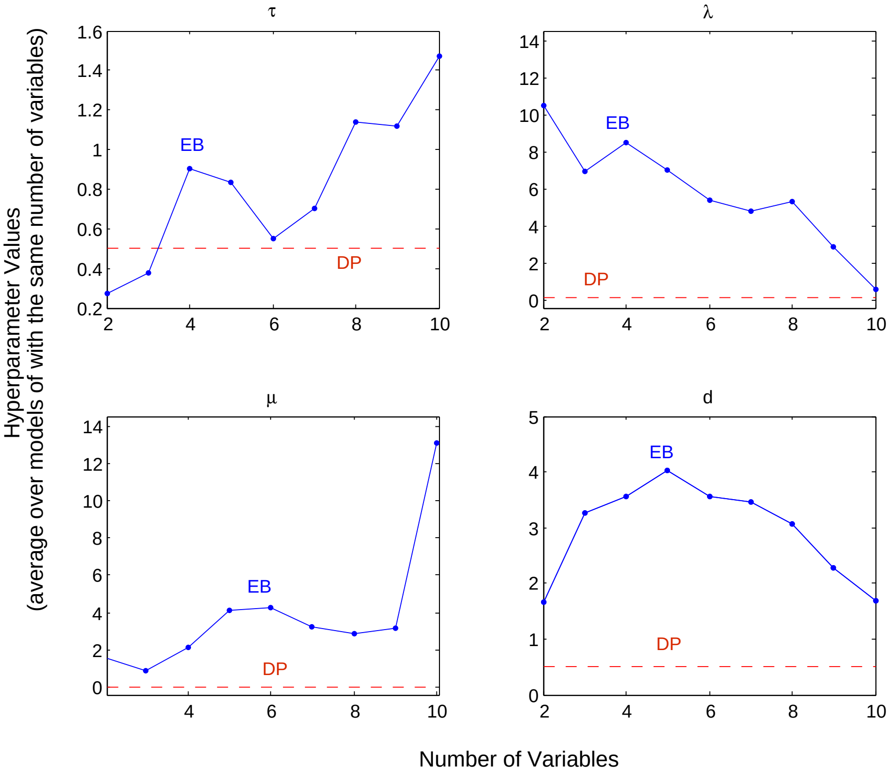
$H0 : L(E_t^2) = L(Erw_t^2)$

$H1 : L(E_t^2) \neq L(Erw_t^2)$

E_t is our nowcast for GDP growth at time t minus the actual growth figure for quarter t. The error associated to the random walk model is defined by Erw_t .

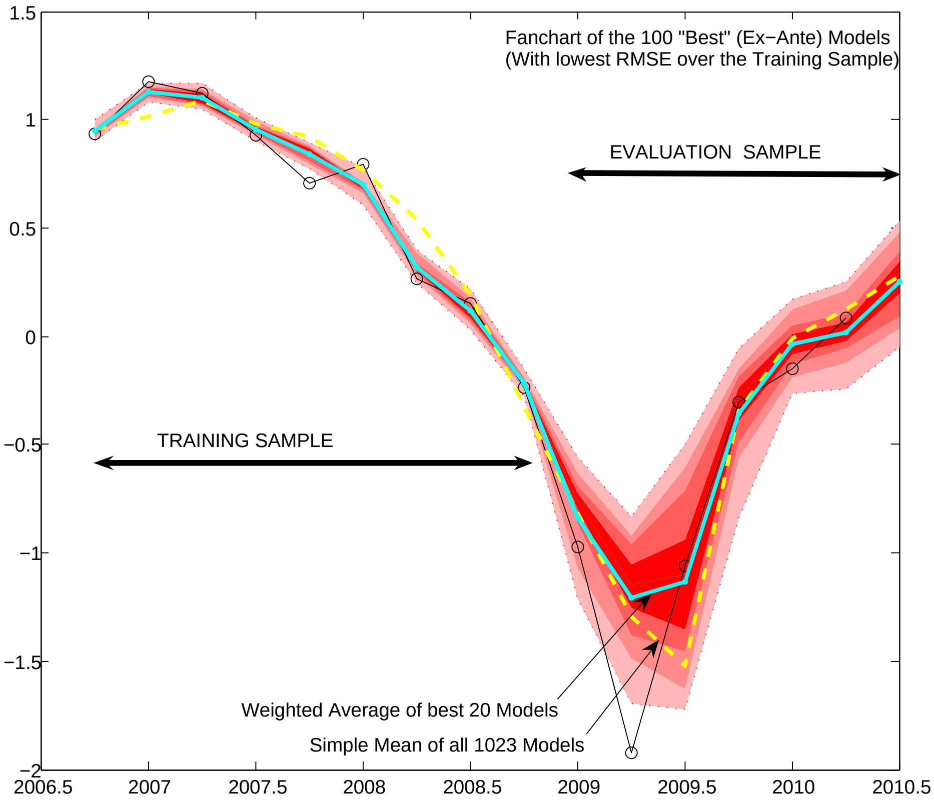
Even if the test has very low power when the sample size is small, i.e. the probability to reject the null of equal forecast accuracy is low even if the null is false, we strongly reject the hypothesis of equal predictive accuracy in the second and third subsamples whenever all models are combined. For the first subsample the results are not favorable simply because the forecast combination does not use hyper-parameterized regressions based on the emprirical Empirical Bayes method. The reason is that we do not have a pre-sample to define the prior.

Figure 1: The Tightness of the Prior



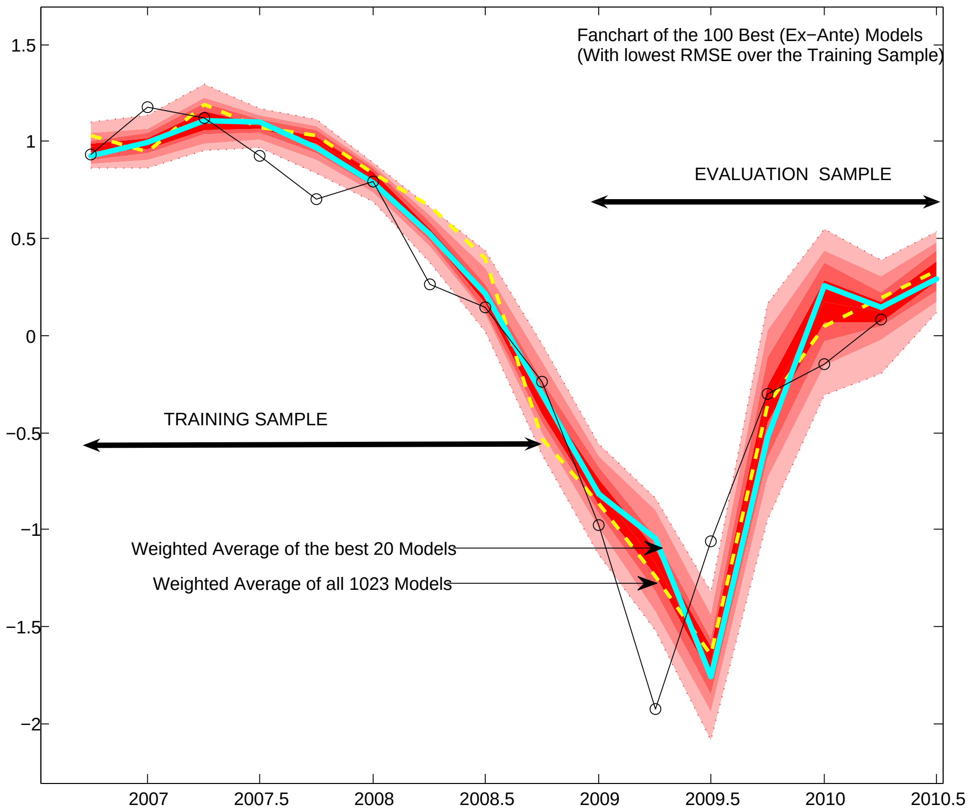
The figures display the average value of the hyperparameters estimated with the EB approach for models with the same number of predictor variables. The number of models of size equal to 2, 3, 4, 5, 6, 7, 8, 9 and 10 is equal to 10, 45, 120, 210, 252, 210, 120, 45 and 10, respectively. The precise definition 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.

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



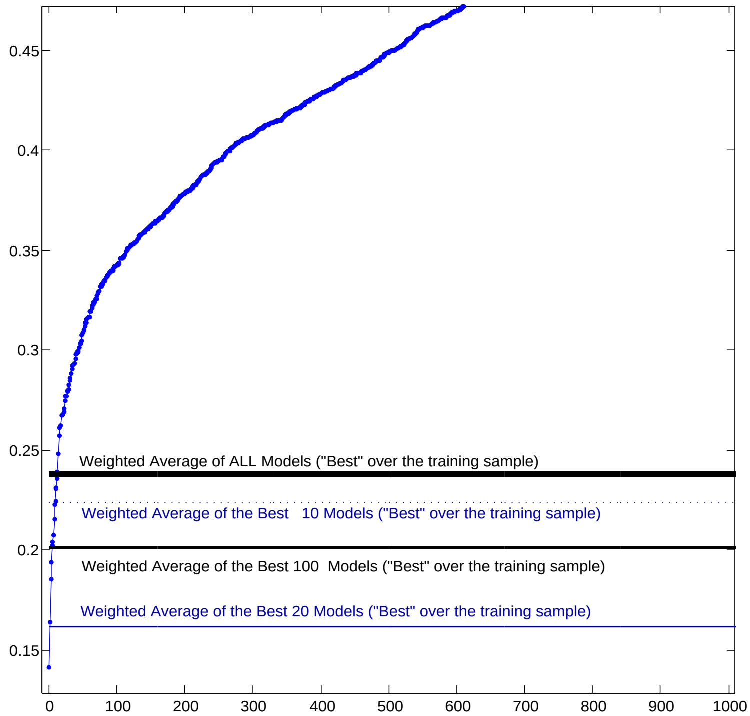
The black circles represent real GDP growth as initially published by the statistical agency. Given that we use the training sample to form priors, it is not surprising that the 10% best performing models provide a perfect fit for GDP growth. The question of interest is whether those models "selected" 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)



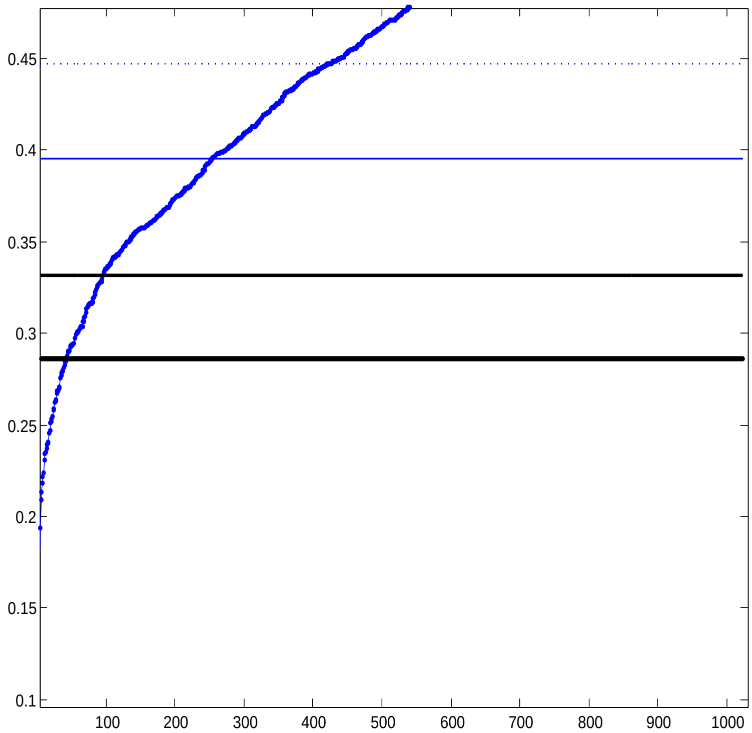
The black circles represent real GDP growth as initially published by the statistical agency. The thick line represents the weighted average nowcast of the 20 models with smallest RMSE over the first subsample. The question of interest is whether those models “selected” on the basis of their performance 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)



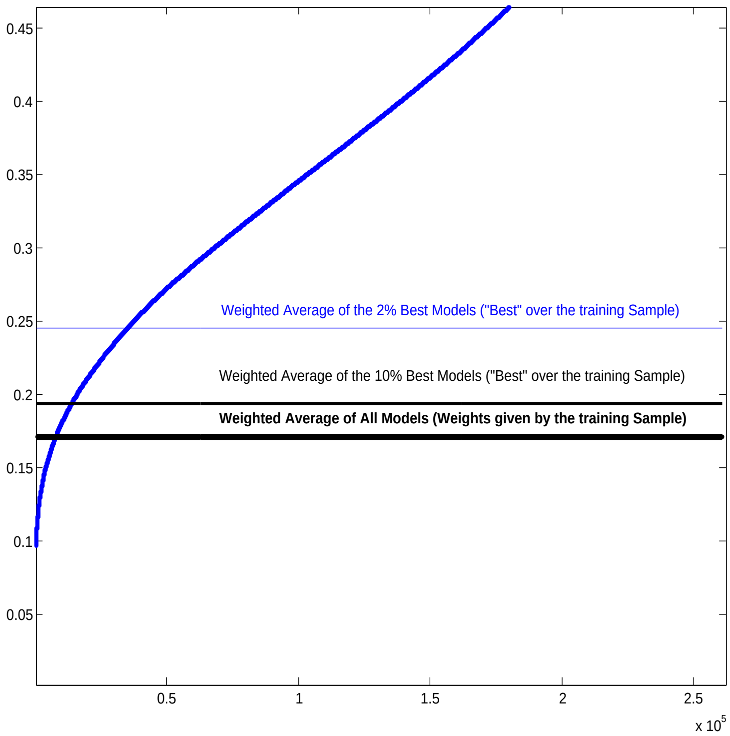
The Root-Mean-Squared Error (RMSE) for each model is computed on the basis of real-time out-of-sample forecast errors for GDP growth. The prediction error is defined as the difference between the nowcast and the last available GDP growth release published by the statistical agency. The RMSEs of all models 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)



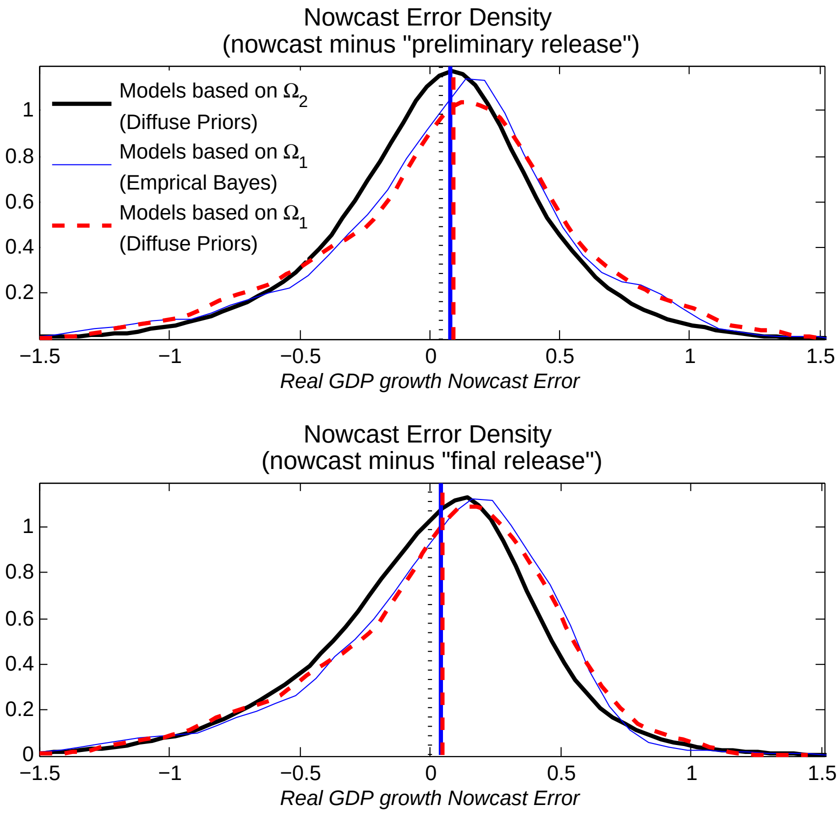
The Root-Mean-Squared Error (RMSE) for each model is computed on the basis of real-time out-of-sample forecast errors for GDP growth. The prediction error is defined as the difference between the nowcast and the last available GDP growth release published by the statistical agency. The RMSEs of all models 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)



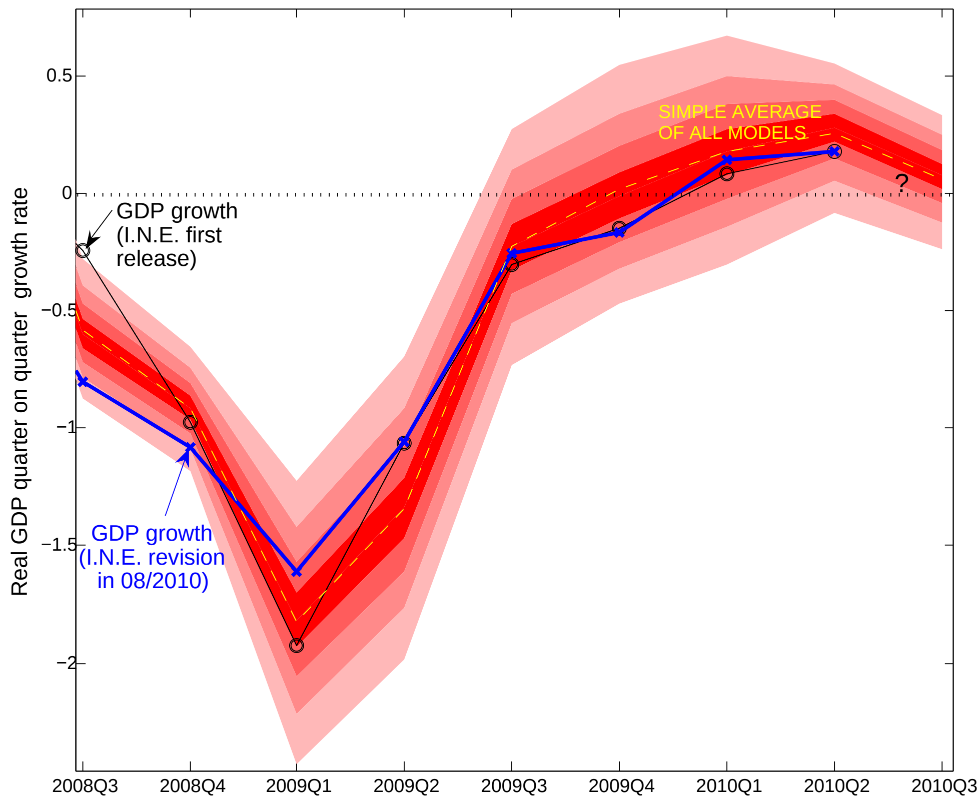
The Root-Mean-Squared Error (RMSE) for each model is computed on the basis of real-time out-of-sample forecast errors for GDP growth. The prediction error is defined as the difference between the nowcast and the last available GDP growth release published by the statistical agency. The RMSEs of all models 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)



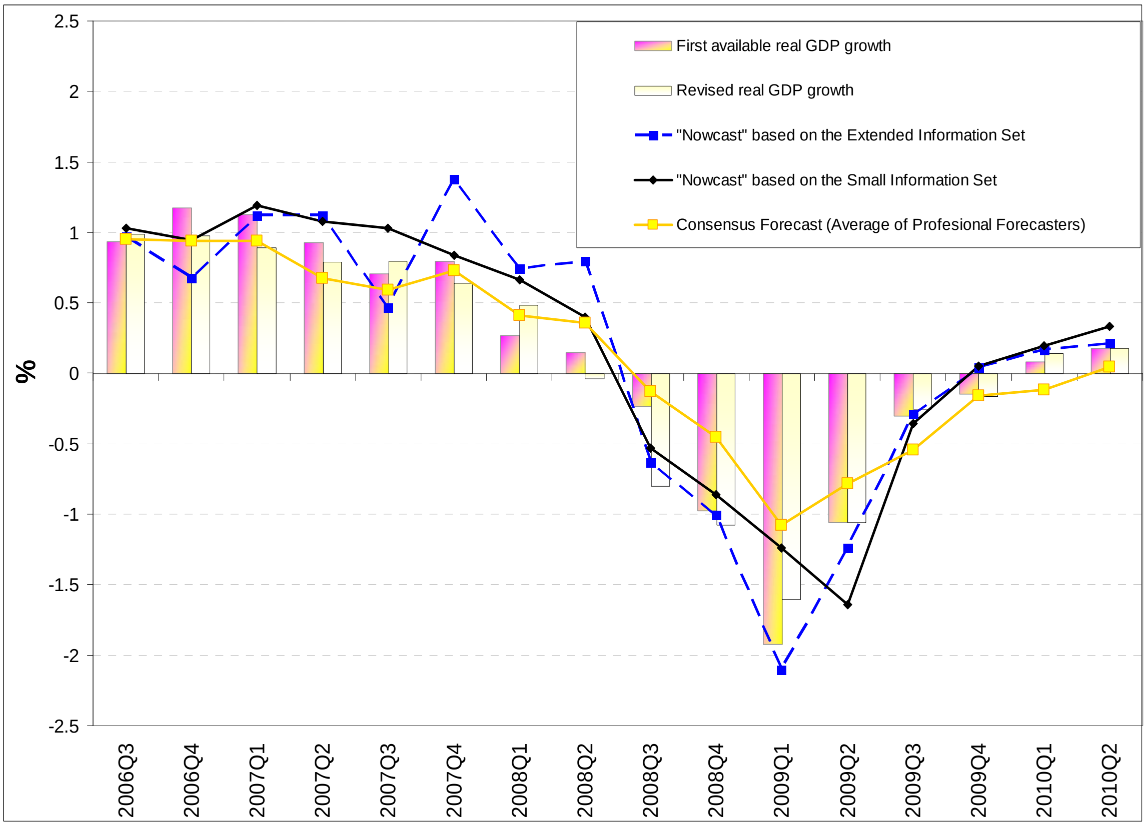
All the projection models obtained under Ω_1 and Ω_2 yield thousands of time series of forecast errors corresponding to our evaluation sample (2008Q4-2010Q2). These graphs represent the probability distributions of all these forecast errors. The upper figure shows that, when the small information set (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 over prediction 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 necessarily coincide with the mode.

Figure 8: Nowcasts conditional on Ω^2 (Diffuse Prior)



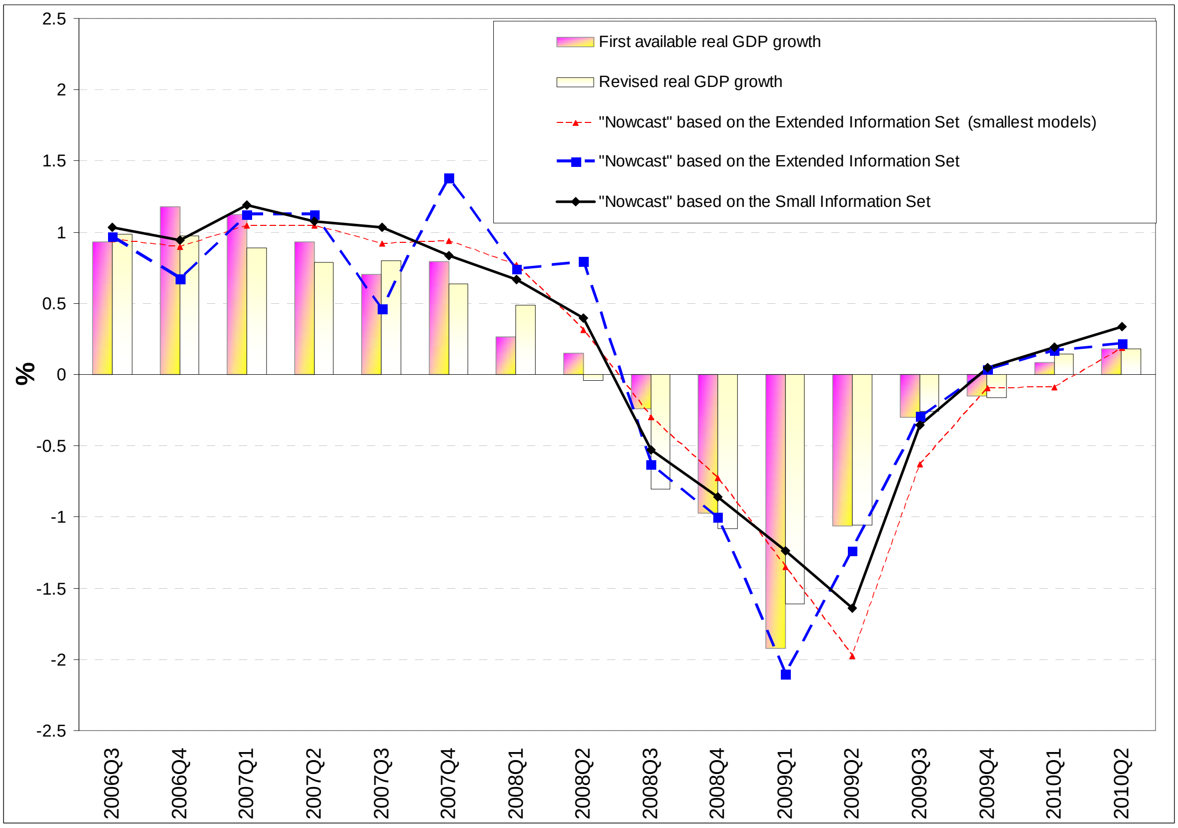
The black circles represent real GDP growth as initially published by the statistical agency. The fan chart represents a 90% forecasting interval that takes into account model uncertainty. The dashed line is a simple average of all 262144 models. This graph also represents the projection exercise for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

Figure 9: Comparison with the "Consensus Forecast"



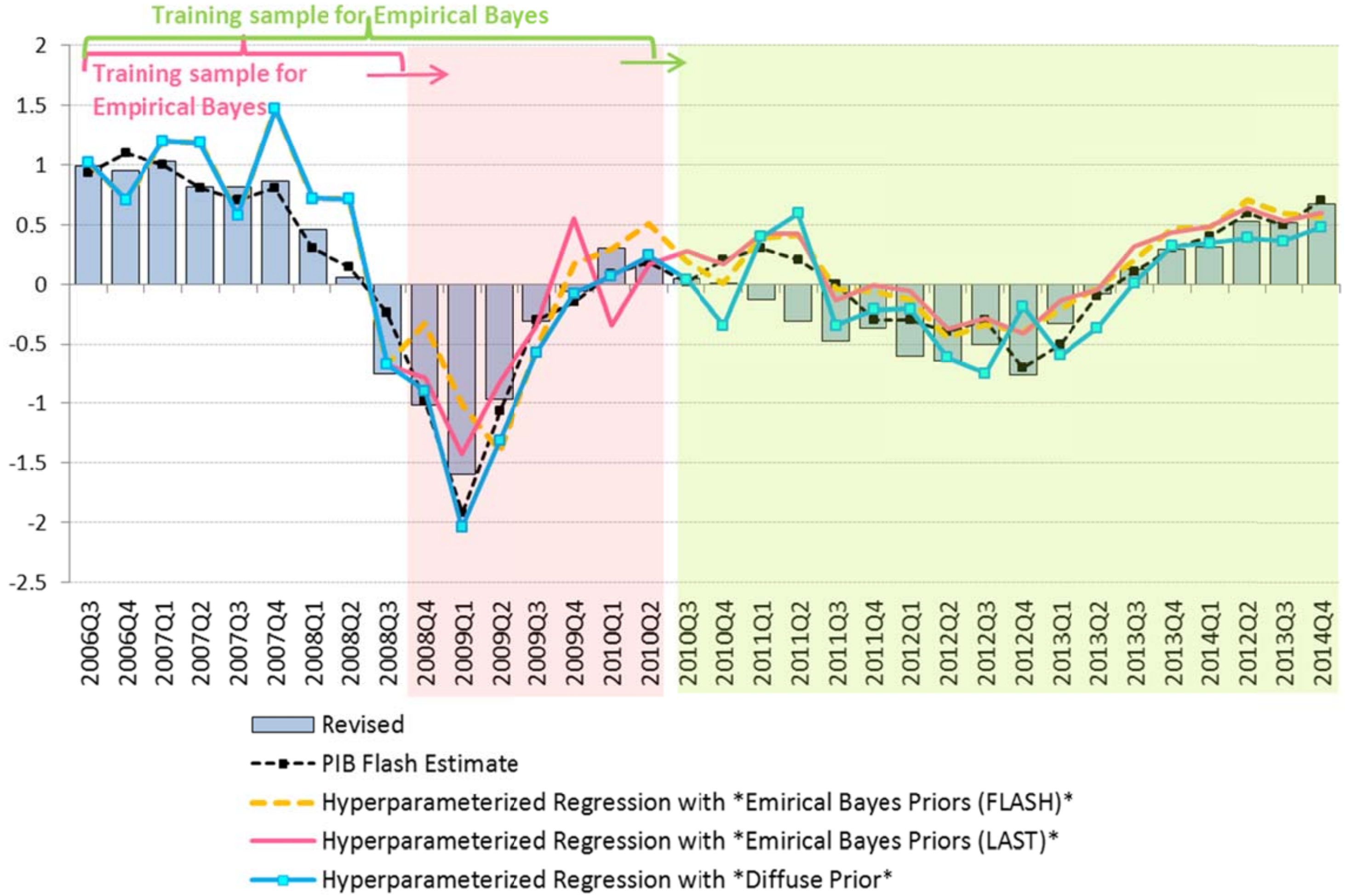
The figure illustrates the forecasting ability of the mean of the survey of professional forecasters compiled by Consensus Economics and published in their monthly publication "Consensus Forecast". This comparison is quite meaningful, since it is also an aggregation of individual forecasts. Moreover, 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 emphasising 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: *Small* models based on the extended information set Ω_2



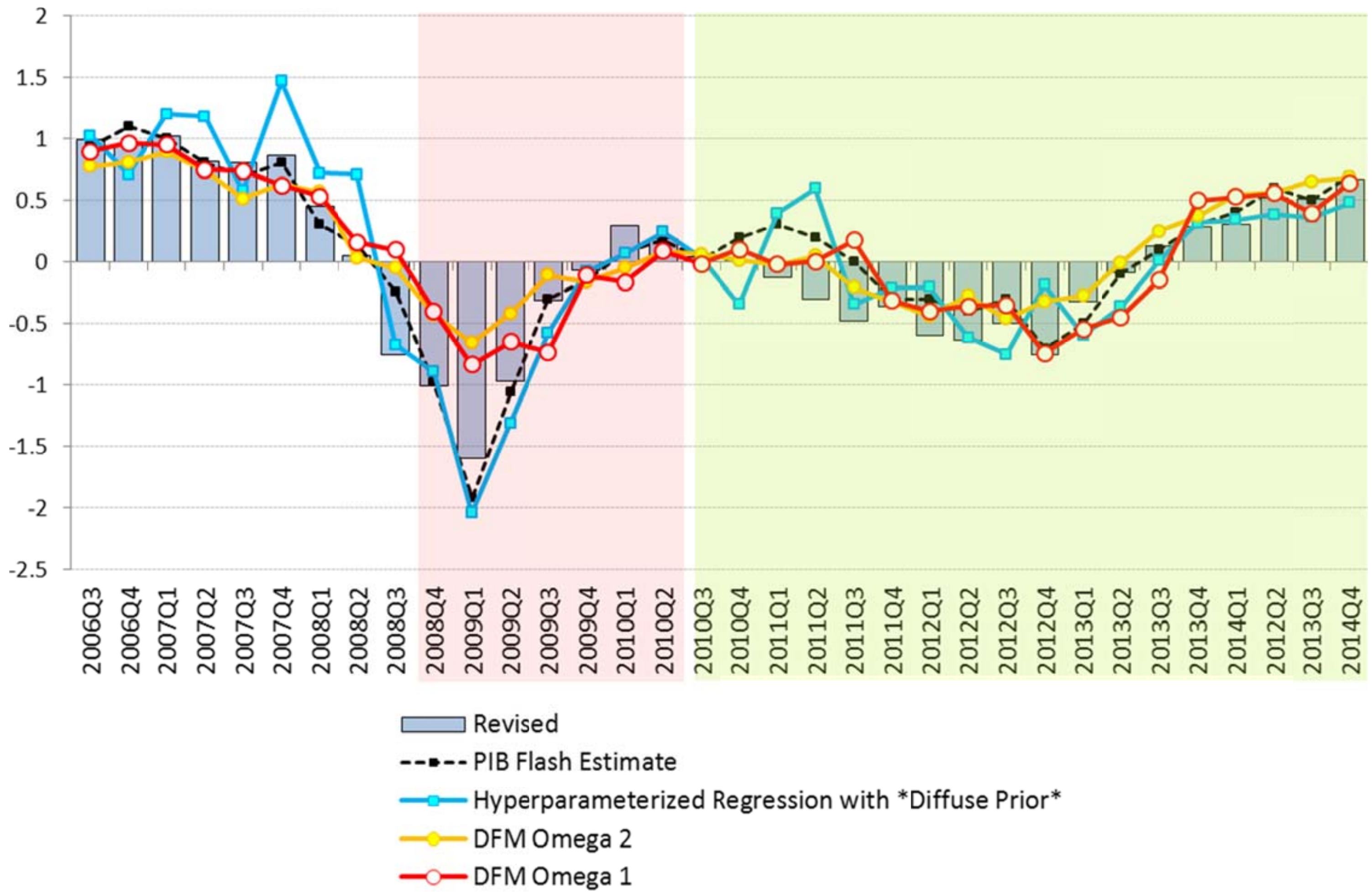
The figure compares the nowcasting performance of a simple average of *large* models based on the set Ω_2 (dashed line with squares) with the one based on the smaller information set Ω_1, which only contains 10 economic indicators other than GDP (solid line). Moreover, we show that a combination of all (*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 11: Nowcasting real GDP growth with Dynamic Factor Models



The Empirical Bayes method for prior elicitation exploits the first subsample for estimation and forecasting during the second subsample. It also uses the second subsample to elicit the priors for estimation and forecasting during the third subsample. In the absence of a training sample prior to 2006$Q3$, all forecasts for the first subsample use the so-called diffuse prior, which produces practically the same results as with OLS.

Figure 12: Nowcasting real GDP growth with Hyper-Parameterized Bayesian Regressions



The end of quarter nowcasts coming from two identically specified dynamic factor models based on Ω_1 (Small) and Ω_2 (Large) respectively, are compared in this graph to those resulting from the hyper-parameterized dynamic regression model estimated with diffuse priors. That regression model incorporates all indicators in Ω_2 in their lags.