MODEL AVERAGING IN MARKOV-SWITCHING MODELS: PREDICTING NATIONAL RECESSIONS WITH REGIONAL DATA

2017

Pierre Guérin and Danilo Leiva-Leon

Documentos de Trabajo
N.º 1727

BANCO DE **ESPAÑA**
Eurosistema

MODEL AVERAGING IN MARKOV-SWITCHING MODELS:
PREDICTING NATIONAL RECESSIONS WITH REGIONAL DATA

MODEL AVERAGING IN MARKOV-SWITCHING MODELS: PREDICTING NATIONAL RECESSIONS WITH REGIONAL DATA (*)

Pierre Guérin (**)

BANK OF CANADA

Danilo Leiva-Leon (***)

BANCO DE ESPAÑA

(*) We would like to thank Massimiliano Marcellino, Lutz Kilian, Michael Owyang, James Hamilton, Ana Galvão, Jeremy Piger, Robert Lavigne, Michael Ehrmann and seminar participants at the Applied Time Series Econometrics Workshop held at the Federal Reserve Bank of St. Louis, the 2014 CFE conference in Pisa, Bank of Canada, University of California at Riverside, and the 2015 Canadian Econometric Study Group for helpful discussions and comments. An earlier version of this paper was circulated as *Bank of Canada Working Paper* No. 2015-24. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada or the Bank of Spain.
(**) Bank of Canada, e-mail: pjg.guerin@gmail.com
(***) Bank of Spain, e-mail: danilo.leiva@bde.es

Documentos de Trabajo. N.º 1727

2017

ISSN: 1579-8666 (on line)

Abstract

This paper introduces new weighting schemes for model averaging when one is interested in combining discrete forecasts from competing Markov-switching models. In the empirical application, we forecast U.S. business cycle turning points with statelevel employment data. We nd that forecasts obtained with our best combination scheme provide timely updates of U.S. recessions in that they outperform a notoriously dicult benchmark to beat (the anxious index from the Survey of Professional Forecasters) for short-term forecasts.

Keywords: business cycles, forecast combination, forecasting, Markov-switching, nowcasting.

JEL classification: C53, E32, E37.

Resumen

En este artículo se introducen nuevos esquemas de ponderación para promediar de modelos econométricos cuando se está interesado en combinar predicciones de variables discretas provenientes de modelos con cambios de régimen markoviano. En una aplicación empírica, se pronostican los puntos de inflexión de los ciclos económicos de Estados Unidos con datos de empleo a nivel de estados. Los pronósticos obtenidos con los esquemas de ponderación propuestos proporcionan actualizaciones oportunas de las recesiones en Estados Unidos.

Palabras clave: ciclos económicos, combinación de predicciones, cambios markovianos.

Códigos JEL: C53, E32, E37.

1 Introduction

Combining forecasts has frequently been found to produce better forecasts on average than forecasts obtained from the best ex-ante individual forecasting model. While there is a large literature on model averaging for linear models, little work has been done to study model-averaging schemes for Markov-switching models.[1] In this paper, we introduce new weighting schemes to combine discrete forecasts from competing Markov-switching models. We show their relevance based on an empirical application to predict U.S. recessions with a large set of Markov-switching models.[2]

In the context of linear regressions, an approach increasingly used in empirical studies is Bayesian model averaging (BMA), proposed by Raftery et al. (1998). This approach is often referred to as static model averaging in that the estimated models' weights are constant over time. Recently, Raftery et al. (2010) proposed the dynamic model averaging (DMA) approach, where the models' weights evolve over time.

This paper contributes to the literature along two dimensions. First, we introduce an algorithm to perform DMA with Markov-switching models. This extension provides a flexible framework that evaluates, at every period of time, the performance of different Markov-switching models to infer the regimes of a target variable. Second, we introduce new weighting schemes for model averaging when the variable to forecast is a discrete outcome. Specifically, we propose models' weights that depend on the past predictive ability of a given model to estimate discrete outcomes. It is intuitive to do so in that a model that performs well for continuous forecasts may not work well for discrete forecasts. Hence, standard weighting schemes that exclusively rely on the likelihood as a measure of model fit may not be appropriate when forecasting discrete outcomes. When presenting our new models' weights, we consider two classes of combination schemes, constant weights (BMA) and time-varying weights (DMA).

Our main results can be summarized as follows. First, in our empirical experiment, we find that it is relevant to take into account the models' ability to estimate regimes when calculating models' weights if one is interested in regime classification. Indeed, our combination schemes based on the predictive ability to fit discrete outcomes typically outperform combination schemes based on the likelihood only. This holds true in an out-of-sample forecasting experiment. Second, on average, the best combination scheme in terms of predictive accuracy is obtained with the DMA framework where the (time-varying) weights depend on the past predictive ability to estimate discrete outcomes. Third, the use of regional data improves the forecasting performance compared with models using exclusively national data. Fourth, out-of-sample forecasts obtained with the best combination scheme outperform the anxious index from the Survey of Professional Forecasters for short-term forecasts, which shows the empirical relevance of our framework.

[1] Among the notable exceptions are Billio et al. (2012), who compare the performance of combination schemes for linear and regime-switching models, and Billio et al. (2013), who propose a time-varying combination approach for multivariate predictive densities.

[2] The working paper version of this article also reports Monte Carlo evidence to evaluate the relevance of our proposed combination schemes in a controlled experiment.

2 Combination schemes

We consider a parsimonious regime-switching regression model defined as follows:

$$y_t = \mu^k_{S_t} + \beta^k x^k_t + v^k_t, \tag{1}$$

where y_t is the dependent variable and x^k_t denotes a given regressor k. The error term, denoted by v^k_t, is assumed to be normally distributed, that is, $v^k_t \sim N(0, \sigma^2_k)$. The regime-switching intercept is given by $\mu^k_{S_t} = (\mu^k_0 + \mu^k_1 S^k_t)$, where S^k_t is a Markov chain defined by the following constant transition probability:

$$p^k_{ij} = P(S^k_{t+1} = j | S^k_t = i), \tag{2}$$

$$\sum_{j=1}^{2} p^k_{ij} = 1 \ \ \forall i, j \epsilon \{1, 2\}. \tag{3}$$

Note also that equation (1) could easily accommodate multiple regressors instead of a single predictor x^k_t as well as regime-switching parameters for the slope coefficients and error variance. Moreover, for ease of exposition, we assume that S_t is only governed by two regimes, but the framework we develop in this paper can easily accommodate more than two states.

2.1 Bayesian model averaging

Suppose that we have K different models, M_k for $k = 1, 2, ..., K$, which all seek to explain y_t, according to equation (1). If one is interested in comparing different models, we can use Bayes' rule to derive the posterior model probability and assess the degree of support for model k,

$$f(M_k | y_t) = \frac{f(y_t | M_k) f(M_k)}{f(y_t)}, \tag{4}$$

where $f(y_t) = \sum_{j=1}^{K} f(y_t | M_j) f(M_j)$, $f(M_k)$ is the prior probability that model k is true and $f(y_t | M_k)$ is the marginal likelihood for model k. Accordingly, we define $f(M_k | y_t)$ in equation (4), for $k = 1, 2, ..., K$, as the *likelihood-based* static weighting scheme.[3]

Given that our goal is to predict the discrete variable S_t rather than the continuous variable y_t, we propose an alternative weighting scheme to reflect this objective. In doing so, we use Bayes' rule to derive a probability statement about the most appropriate model M_k to explain the regimes S_t conditional on the data and the estimated probability of being in a given regime

$$\begin{aligned} f(M_k | y_t, S_t) &= \frac{f(y_t, S_t | M_k) f(M_k)}{f(y_t, S_t)} && (5) \\ &= \frac{f(S_t | y_t, M_k) f(y_t | M_k) f(M_k)}{\sum_{j=1}^{K} f(S_t | y_t, M_j) f(y_t | M_j) f(M_j)}, && (6) \end{aligned}$$

[3]Given that the parameters are estimated with the Gibbs sampler, to improve efficiency, we follow Newton and Raftery (1994) in computing the marginal likelihood and use the harmonic mean estimator, which is a simulation-consistent estimate that uses samples from the posterior density.

where the term $f(S_t|y_t, M_k)$ indicates the model's ability to fit S_t. We propose to use the inverse quadratic probability score (QPS) to evaluate $f(S_t|y_t, M_k)$, since it is a measure commonly used to evaluate discrete outcomes.[4] The QPS associated with model k is defined as follows

$$QPS_k = \frac{2}{T}\sum_{t=1}^{T}(P(S_t^k = 1|\psi_t) - S_t)^2, \tag{7}$$

where the lower the QPS, the better the ability of the model to fit S_t. Therefore, the posterior model probability for model k reads as

$$f(M_k|y_t, S_t) = \frac{f(y_t|M_k)f(M_k)QPS_k^{-1}}{\sum_{j=1}^{K} f(y_t|M_j)f(M_j)QPS_j^{-1}}. \tag{8}$$

Accordingly, we define $f(M_k|y_t, S_t)$ in equation (8), for $k = 1, 2, ..., K$, as the *combination-based* static weighting scheme, since it combines goodness-of-fit criteria for both y_t and S_t.

However, since we are only interested in assessing the ability of model M_k to explain S_t, we also could avoid conditioning on y_t and propose the following expression for the posterior probability model

$$\begin{aligned} f(M_k|S_t) &= \frac{f(S_t|M_k)f(M_k)}{\sum_{j=1}^{K} f(S_t|M_j)f(M_j)} & (9)\\ &= \frac{f(M_k)QPS_k^{-1}}{\sum_{j=1}^{K} f(M_j)QPS_j^{-1}}. & (10)\end{aligned}$$

Accordingly, we define $f(M_k|S_t)$ in equation (9), for $k = 1, 2, ..., K$, as the *QPS-based* static weighting scheme.

2.2 Dynamic model averaging

Dynamic model averaging (DMA) has first been implemented in econometrics by Koop and Korobilis (2012) to perform model averaging for time-varying parameter regression where the weights also vary over time. In our context, to calculate the time-varying weights for each of the Markov-switching models we estimate, we introduce the following algorithm that combines the Hamilton filter with the prediction and updating equations used in the DMA approach from Raftery et al. (2010).

Suppose that we have K different models, and let M_t be the indicator variable for the models at time t, that is, $M_t = \{1, 2, ..., K\}$. At any given period t, we compute the following steps for all the K models under consideration:

[4]Note that alternative criteria could be used to evaluate the models' ability to classify regimes. For example, the logarithmic probability score or the area under the receiver operating characteristics (see, e.g., Berge and Jordà (2011)) could be used.

Step 1: Compute the predicted regime probabilities for any given model k given past information ψ_{t-1} as follows

$$P(S_t^k = j|\psi_{t-1}) = \sum_{S_{t-1}^k} P(S_t^k = j, S_{t-1}^k = i|\psi_{t-1}), \tag{11}$$

where $P(S_t^k = j, S_{t-1}^k = i|\psi_{t-1}) = P(S_t^k = j|S_{t-1}^k = i)P(S_{t-1}^k = i|\psi_{t-1})$, and $P(S_t^k = j|S_{t-1}^k = i)$ are the corresponding transition probabilities. Initial regime probabilities $P(S_0|\psi_0)$ can be obtained by using the ergodic probabilities.

Step 2: Compute the predicted probability associated with the k-th model by using the forgetting factor α as in Raftery et al. (2010)

$$P(M_t = k|\psi_{t-1}) = \frac{P(M_{t-1} = k|\psi_{t-1})^\alpha}{\sum_{M_{t-1}} P(M_{t-1} = j|\psi_{t-1})^\alpha}. \tag{12}$$

The forgetting factor α is the coefficient that governs the amount of persistence in the models' weights, and it is set to a fixed value slightly less than one. The higher the α, the higher the weight attached to past predictive performance is.[5] Initial model probabilities $P(M_0|\psi_0)$ can be obtained using equal weights for all models.

Step 3: As new information arrives, y_t, compute the updated regime probabilities for any given model k as follows

$$P(S_t^k = j|\psi_t) = \sum_{S_{t-1}^k} P(S_t^k = j, S_{t-1}^k = i|\psi_t), \tag{13}$$

where

$$P(S_t^k = j, S_{t-1}^k = i|\psi_t) = \frac{f_k(y_t|S_t^k = j, S_{t-1}^k = i, \psi_{t-1})P(S_t^k = j, S_{t-1}^k = i|\psi_{t-1})}{f_k(y_t|\psi_{t-1})}. \tag{14}$$

The term $f_k(y_t|S_t^k = j, S_{t-1}^k = i, \psi_{t-1})$ is the conditional likelihood from the corresponding model, the term $P(S_t^k = j, S_{t-1}^k = i|\psi_{t-1})$ is obtained from Step 1, and the predictive likelihood is given by

$$f_k(y_t|\psi_{t-1}) = \sum_{S_t^k} \sum_{S_{t-1}^k} f_k(y_t|S_t^k = j, S_{t-1}^k = i, \psi_{t-1})P(S_t^k = j, S_{t-1}^k = i|\psi_{t-1}). \tag{15}$$

Step 4: Compute the updated probability associated to the k-th model following the updating criterion of Raftery et al. (2010), which is based on a measure of model fit for y_t, that is, the predictive likelihood

$$P(M_t = k|\psi_t) = \frac{P(M_t = k|\psi_{t-1})f_k(y_t|\psi_{t-1})}{\sum_{M_t} P(M_t = j|\psi_{t-1})f_j(y_t|\psi_{t-1})}. \tag{16}$$

[5] In the empirical application, we use two different values for α, $\alpha = 0.99$ or $\alpha = 0.95$, implying that forecasting performance from two years ago receives about 78.5 per cent or 29 per cent weight compared with last period's forecasting performance, respectively.

We repeat the four steps above for each model at each period of time $t = 1, ..., T$. Accordingly, we define $P(M_t = k|\psi_t)$ in equation (16), for $k = 1, 2, ..., K$, as the *likelihood-based* dynamic weighting scheme. Also, since the output of the algorithm consists of the regime probabilities for each model and the model probabilities for each time period we can compute the expected regime probabilities by averaging them across models:

$$P(S_t = j|\psi_t) = \sum_{k=1}^{K} P(S_t^k = j|\psi_t)P(M_t = k|\psi_t). \tag{17}$$

The average probability $P(S_t = j|\psi_t)$ is used to assess the performance of the proposed Markov-switching DMA combination scheme.[6]

In line with Section 2.1, we also allow for the possibility that both the marginal likelihood and the QPS criterion could indicate the model's ability to predict S_t, as follows. The updating equation (16) is replaced by

$$P(M_t = k|\psi_t, S_t) = \frac{P(M_t = k|\psi_{t-1})f_k(y_t|\psi_{t-1})Q_{t|t,k}^{-1}}{\sum_{M_t} P(M_t = j|\psi_{t-1})f_j(y_t|\psi_{t-1})Q_{t|t,k}^{-1}}, \tag{18}$$

where $Q_{t|t,k}$ is the cumulative QPS at time t for model k, defined as

$$Q_{t|t,k} = \frac{2}{t}(\sum_{\tau=1}^{t} P(S_\tau^k = 1|\psi_\tau) - S_\tau)^2, \tag{19}$$

and the model prediction equation remains the same as in equation (12). Accordingly, we define $P(M_t = k|\psi_t, S_t)$ in equation (18), for $k = 1, 2, ..., K$, as the *combination-based* dynamic weighting scheme.

Again, since we are interested in predicting S_t rather than y_t, we modify the Raftery et al. (2010) approach and replace the marginal likelihood, which measures how well the model fits y_t, with a measure of goodness-of-fit for S_t, that is

$$P(M_t = k|\psi_t, S_t) = \frac{P(M_t = k|\psi_{t-1})Q_{t|t,k}^{-1}}{\sum_{M_t} P(M_t = j|\psi_{t-1})Q_{t|t,k}^{-1}}, \tag{20}$$

Accordingly, we define $P(M_t = k|\psi_t, S_t)$ in equation (20), for $k = 1, 2, ..., K$, as the *QPS-based* dynamic weighting scheme.

[6]Notice that DMA also differs from BMA in that no simulation is required to calculate the weights.

3 Empirical Results

3.1 Data

We use alternatively industrial production and employment data as a measure of national economic activity. These two indicators are available on a monthly basis, and are important measures of economic activity in the United States. The state-level data we use are the employees on non-farm payrolls data series published at a monthly frequency for all 50 U.S. states by the U.S. Bureau of Labor Statistics (note that state-level data are not available for monthly industrial production). These state-level data are available on a not seasonally adjusted basis since at least January 1960 for all U.S. states, and seasonally adjusted data are available since January 1990. We use real-time data for the nationwide dependent variable and for the NBER recession dummy variable, and revised series for the state-level data.[7] All data are taken as 100 times the change in the log-level of the series to obtain monthly percent changes. To facilitate inference on the regimes, and obtain a long enough evaluation sample to assess the accuracy of the forecasts, we use data starting from 1960, and the state-level data are appropriately seasonally adjusted using the X-12 seasonal adjustment methodology. Hence, the full estimation sample extends from February 1960 to April 2014.[8]

3.2 Out-of-sample results

The first estimation sample extends from February 1960 to December 1978, and it is recursively expanded until September 2013, that is, the evaluation sample covers the period ranging from January 1979 to March 2014 (i.e., the last forecast six months ahead refers to the month of March 2014). As such, our evaluation sample includes five recessions that cover 13.2 per cent of the sample, which permits us to mitigate the risks of spurious forecasting results. The models are re-estimated every month as new information becomes available.

We formulate forecasts for horizon $h = \{0, 1, 2, 3, 6\}$, that is, from the current month ($h = 0$) up to six months ahead ($h = 6$). We use the quadratic probability score (QPS) to evaluate the accuracy in predicting turning points. The out-of-sample QPS (QPS^{OOS}) is defined as follows:

$$QPS_k^{OOS} = \frac{2}{T - T_0 + 1} \sum_{t=T_0}^{T} (P(S_{t+h}^k = 0|\psi_t) - NBER_{t+h})^2, \tag{21}$$

[7] Real-time data on state-level employment are available only since June 2007 from the "Alfred" real-time database of the Federal Reserve Bank of St. Louis (`http://alfred.stlouisfed.org/`). This data constraint precludes us to perform a fully real-time out-of-sample forecasting exercise over a long enough evalutation sample. However, in the working paper version of this article, we analyze the performance of the weighting schemes in a fully real-time setting using data since June 2007 to examine the predictive performance of the different combinaton schemes during the last U.S. recession.

[8] We follow the multi-move Gibbs-sampling procedure in Kim and Nelson (1999) to estimate the parameters and produce inference on regimes for the univariate Markov-switching models. Full details on the estimation algorithm and prior distributions are available in chapter 9 of Kim and Nelson (1999) (and in the working paper version of this article). All models are estimated using 7000 draws, discarding the first 2000 draws to mitigate the effect of the initial conditions.

where $T - T_0 + 1$ is the size of the evaluation sample, $P(S^k_{t+h} = 0|\psi_t)$ is the probability of being in the first regime (i.e., the recession regime) in period $t + h$, and $NBER_{t+h}$ is a dummy variable that takes on a value of 1 if the U.S. economy is in recession in period $t+h$ according to the NBER business cycle dating committee and 0 otherwise. The predicted probabilities of being in regime j from model k, $P(S^k_{t+h} = j|\psi_t)$, are calculated as follows:

$$P(S^k_{t+h} = j|\psi_t) = \sum_{i=1}^{M} p^k_{ij} P(S^k_{t+h-1} = j|\psi_t), \tag{22}$$

where M denotes the maximum number of regimes (two in this case) and p^k_{ij} is the transition probability of going from regime i to regime j from model k (i.e., $p^k_{ij} = P(S^k_{t+1} = j|S^k_t = i)$, calculated as the median of the parameter estimates over the 5000 simulations performed to calculate the posterior distributions of these parameters. The predicted probabilities from each model are then averaged at each point in time of the evaluation sample using the combination schemes outlined in section 2.

In comparing models, we also report results obtained from using the anxious index from the Survey of Professional Forecasters (SPF) of the Philadelphia Federal Reserve Bank. This index corresponds to the probability of a decline in real GDP. This is a relevant benchmark, since survey forecasts have been found to perform well compared with model-based predictions (see, e.g., Faust and Wright (2009)). The SPF is available only on a quarterly basis, but we disaggregate it at the monthly frequency assuming that its monthly value is constant over the three months of the quarter. Moreover, we also evaluate the statistical significance of our results using the Diebold-Mariano-West test to assess equal out-of-sample predictive accuracy (see Diebold and Mariano (1995) and West (1996)), using the likelihood-based weighting scheme as a benchmark model. In this way, we evaluate from a statistical point of view the relevance of our weighting scheme based on the QPS compared with the traditional approach that relies exclusively on the likelihood.

Table 1 reports the results. First, the combination scheme with industrial production using DMA weights based on the QPS obtains the best forecasting results for forecast horizons $h = \{0, 1\}$, and the SPF anxious index obtained the best results for forecast horizons $h = \{2, 3, 6\}$. Second, the QPS-based combination schemes nearly always outperform the combination schemes based on the likelihood only, and typically in a statistically significant way. Overall, our results show the importance of state-level data to estimate and forecast U.S. business cycle turning points. Moreover, our results also emphasize the relevance of forecast combination based on time-varying weights, which are obtained from the ability of each model to estimate a discrete outcome.

4 Conclusions

In this letter, we first introduce an algorithm to use dynamic model averaging to combine forecasts from a large set of Markov-switching models. Second, we modify the standard Bayesian model averaging (BMA) and dynamic model averaging (DMA) combination schemes to make the weights depend on past performance in order to detect regime changes using the quadratic probability score (QPS) to measure the models' ability to classify regimes. Third, our empirical results show that our proposed combination schemes outperform competing combination schemes in terms of forecasting accuracy to predict U.S. recessions. Therefore standard weighting schemes based only on the models' likelihood are not necessarily appropriate in a context of regime classification.

References

Berge, T. J. and Jordà, O. (2011). Evaluating the Classification of Economic Activity into Recessions and Expansions. *American Economic Journal: Macroeconomics*, 3(2):246–77.

Billio, M., Casarin, R., Ravazzolo, F., and van Dijk, H. K. (2012). Combination schemes for turning point predictions. *The Quarterly Review of Economics and Finance*, 52(4):402–412.

Billio, M., Casarin, R., Ravazzolo, F., and van Dijk, H. K. (2013). Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics*, 177(2):213–232.

Diebold, F. X. and Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 13(3):253–63.

Faust, J. and Wright, J. H. (2009). Comparing Greenbook and Reduced Form Forecasts Using a Large Realtime Dataset. *Journal of Business & Economic Statistics*, 27(4):468–479.

Kim, C.-J. and Nelson, C. R. (1999). *State-space models with regime switching: classical and Gibbs-sampling approaches with applications.* The MIT Press.

Koop, G. and Korobilis, D. (2012). Forecasting Inflation Using Dynamic Model Averaging. *International Economic Review*, 53(3):867–886.

Newton, M. A. and Raftery, A. E. (1994). Approximate Bayesian inference with the weighted likelihood boostrap. *Journal of the Royal Statistical Society: Series B*, 56:3–26.

Raftery, A., Karny, M., and Ettler, P. (2010). Online Predictions Under Model Uncertainty Via Dynamic Model Averaging: Application to a Cold Rolling Mill. *Technometrics*, 52:52–66.

Raftery, A. E., Madigan, D., and Hoeting, J. A. (1998). Bayesian Model Averaging for Linear Regression Models. *Journal of the American Statistical Association*, 92:179–191.

West, K. D. (1996). Asymptotic Inference about Predictive Ability. *Econometrica*, 64(5):1067–84.

Table 1: Out-of-sample Quadratic Probability Score

	Employment					
	Forecast horizon (months)	0	1	2	3	6
Dynamic Model	Likelihood-based	0.409	0.388	0.375	0.360	0.329
Averaging	QPS-based	0.208***	0.218***	0.235***	0.247***	0.259***
$\alpha = 0.99$	Combined	0.218***	0.229***	0.246***	0.259***	0.270**
Dynamic Model	Likelihood-based	0.258	0.267	0.275	0.276	0.274
Averaging	QPS-based	0.206	0.222	0.240	0.251	0.266
$\alpha = 0.95$	Combined	0.212	0.229	0.246	0.256	0.265
Bayesian Model	Likelihood-based	0.457	0.429	0.410	0.390	0.350
Averaging	QPS-based	0.223***	0.240***	0.256***	0.264***	0.273***
	Combined	0.457	0.429	0.410	0.390	0.350
Equal weight		0.225***	0.241***	0.257***	0.267***	0.276***
	Industrial Production					
Dynamic Model	Likelihood-based	0.216	0.223	0.233	0.237	0.243
Averaging	QPS-based	**0.133****	**0.161****	0.186**	0.206*	0.235
$\alpha = 0.99$	Combined	0.138**	0.165**	0.189**	0.208*	0.235
Dynamic Model	Likelihood-based	0.189	0.202	0.217	0.226	0.239
Averaging	QPS-based	0.134**	0.162**	0.188*	0.208	0.236
$\alpha = 0.95$	Combined	0.141**	0.166**	0.190**	0.208	0.236
Bayesian Model	Likelihood-based	0.224	0.230	0.238	0.241	0.247
Averaging	QPS-based	0.163**	0.185**	0.206**	0.220**	0.237
	Combined	0.223	0.230	0.238	0.241	0.246
Equal weight		0.153**	0.177**	0.200**	0.216*	0.236
SPF Anxious Index		0.141	0.161	**0.180**	**0.186**	**0.226**
MS-AR (Employment)		0.227	0.238	0.253	0.263	0.278
MS-AR (IP)		0.140	0.166	0.191	0.210	0.236

Note: This table reports the quadratic probability score (QPS) for estimating U.S. business cycle turning points from univariate models using different combination schemes (Bayesian model averaging (BMA), dynamic model averaging (DMA), and an equal-weight scheme for the univariate model described in section 2). The first estimation sample extends from February 1960 to December 1978, and it is recursively expanded until the end of the sample is reached (September 2013). Boldface indicates the model with the lowest QPS for a given horizon. Statistically significant reductions in QPS according to the Diebold-Mariano-West test are marked using ***(1% significance level). **(5% significance level) and *(10% significance level).

BANCO DE ESPAÑA PUBLICATIONS

WORKING PAPERS

1621 ADRIAN VAN RIXTEL, LUNA ROMO GONZÁLEZ and JING YANG: The determinants of long-term debt issuance by European banks: evidence of two crises.

1622 JAVIER ANDRÉS, ÓSCAR ARCE and CARLOS THOMAS: When fiscal consolidation meets private deleveraging.

1623 CARLOS SANZ: The effect of electoral systems on voter turnout: evidence from a natural experiment.

1624 GALO NUÑO and CARLOS THOMAS: Optimal monetary policy with heterogeneous agents.

1625 MARÍA DOLORES GADEA, ANA GÓMEZ-LOSCOS and ANTONIO MONTAÑÉS: Oil price and economic growth: a long story?

1626 PAUL DE GRAUWE and EDDIE GERBA: Stock market cycles and supply side dynamics: two worlds, one vision?

1627 RICARDO GIMENO and EVA ORTEGA: The evolution of inflation expectations in euro area markets.

1628 SUSANA PÁRRAGA RODRÍGUEZ: The dynamic effect of public expenditure shocks in the United States.

1629 SUSANA PÁRRAGA RODRÍGUEZ: The aggregate effects of government incometransfer shocks - EU evidence.

1630 JUAN S. MORA-SANGUINETTI, MARTA MARTÍNEZ-MATUTE and MIGUEL GARCÍA-POSADA: Credit, crisis and contract enforcement: evidence from the Spanish loan market.

1631 PABLO BURRIEL and ALESSANDRO GALESI: Uncovering the heterogeneous effects of ECB unconventional monetary policies across euro area countries.

1632 MAR DELGADO TÉLLEZ, VÍCTOR D. LLEDÓ and JAVIER J. PÉREZ: On the determinants of fiscal non-compliance: an empirical analysis of Spain's regions.

1633 OMAR RACHEDI: Portfolio rebalancing and asset pricing with heterogeneous inattention.

1634 JUAN DE LUCIO, RAÚL MÍNGUEZ, ASIER MINONDO and FRANCISCO REQUENA: The variation of export prices across and within firms.

1635 JUAN FRANCISCO JIMENO, AITOR LACUESTA, MARTA MARTÍNEZ-MATUTE and ERNESTO VILLANUEVA: Education, labour market experience and cognitive skills: evidence from PIAAC.

1701 JAVIER ANDRÉS, JAVIER J. PÉREZ and JUAN A. ROJAS: Implicit public debt thresholds: an empirical exercise for the case of Spain.

1702 LUIS J. ÁLVAREZ: Business cycle estimation with high-pass and band-pass local polynomial regression.

1703 ENRIQUE MORAL-BENITO, PAUL ALLISON and RICHARD WILLIAMS: Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond.

1704 MIKEL BEDAYO: Creating associations as a substitute for direct bank credit. Evidence from Belgium.

1705 MARÍA DOLORES GADEA-RIVAS, ANA GÓMEZ-LOSCOS and DANILO LEIVA-LEON: The evolution of regional economic interlinkages in Europe.

1706 ESTEBAN GARCÍA-MIRALLES: The crucial role of social welfare criteria for optimal inheritance taxation.

1707 MÓNICA CORREA-LÓPEZ and RAFAEL DOMÉNECH: Service regulations, input prices and export volumes: evidence from a panel of manufacturing firms.

1708 MARÍA DOLORES GADEA, ANA GÓMEZ-LOSCOS and GABRIEL PÉREZ-QUIRÓS: Dissecting US recoveries.

1709 CARLOS SANZ: Direct democracy and government size: evidence from Spain.

1710 HENRIQUE S. BASSO and JAMES COSTAIN: Fiscal delegation in a monetary union: instrument assignment and stabilization properties.

1711 IVÁN KATARYNIUK and JAIME MARTÍNEZ-MARTÍN: TFP growth and commodity prices in emerging economies.

1712 SEBASTIAN GECHERT, CHRISTOPH PAETZ and PALOMA VILLANUEVA: Top-down vs. bottom-up? Reconciling the effects of tax and transfer shocks on output.

1713 KNUT ARE AASTVEIT, FRANCESCO FURLANETTO and FRANCESCA LORIA: Has the Fed responded to house and stock prices? A time-varying analysis.

1714 FÁTIMA HERRANZ GONZÁLEZ and CARMEN MARTÍNEZ-CARRASCAL: The impact of firms' financial position on fixed investment and employment. An analysis for Spain.

1715 SERGIO MAYORDOMO, ANTONIO MORENO, STEVEN ONGENA and MARÍA RODRÍGUEZ-MORENO: "Keeping it personal" or "getting real"? On the drivers and effectiveness of personal versus real loan guarantees.

1716 FRANCESCO FURLANETTO and ØRJAN ROBSTAD: Immigration and the macroeconomy: some new empirical evidence.
1717 ALBERTO FUERTES: Exchange rate regime and external adjustment: an empirical investigation for the U.S.
1718 CRISTINA GUILLAMÓN, ENRIQUE MORAL-BENITO and SERGIO PUENTE: High growth firms in employment and productivity: dynamic interactions and the role of financial constraints.
1719 PAULO SOARES ESTEVES and ELVIRA PRADES: On domestic demand and export performance in the euro area countries: does export concentration matter?
1720 LUIS J. ÁLVAREZ and ANA GÓMEZ-LOSCOS: A menu on output gap estimation methods.
1721 PAULA GIL, FRANCISCO MARTÍ, JAVIER J. PÉREZ, ROBERTO RAMOS and RICHARD MORRIS: The output effects of tax changes: narrative evidence from Spain.
1722 RICARDO GIMENO and ALFREDO IBÁÑEZ: The eurozone (expected) inflation: an option's eyes view.
1723 MIGUEL ANTÓN, SERGIO MAYORDOMO and MARÍA RODRÍGUEZ-MORENO: Dealing with dealers: sovereign CDS comovements.
1724 JOSÉ MANUEL MONTERO: Pricing decisions under financial frictions: evidence from the WDN survey.
1725 MARIO ALLOZA: The impact of taxes on income mobility.
1726 DANILO LEIVA-LEON: Measuring business cycles intra-synchronization in US: a regime-switching interdependence framework.
1727 PIERRE GUÉRIN and DANILO LEIVA-LEON: Model averaging in Markov-Switching models: predicting national recessions with regional data.

BANCO DE ESPAÑA
Eurosistema

Unidad de Servicios Auxiliares
Alcalá, 48 - 28014 Madrid
E-mail: publicaciones@bde.es
www.bde.es