



A Toolkit to Strengthen Government Budget Surveillance*

DIEGO J. PEDREGAL

Universidad de Castilla-La Mancha

JAVIER J. PÉREZ

Banco de España

ANTONIO JESÚS SÁNCHEZ-FUENTES

Universidad Complutense de Madrid & GEN

Received: June, 2014

Accepted: December, 2014

Summary

In this paper we develop a comprehensive short-term fiscal forecasting system, useful for real-time monitoring of government's borrowing requirement in Spain, a country that has been at the center of the recent European sovereign debt crisis, not least because of sizeable failures to meet public deficit targets. The system is made of a suite of models, with different levels of disaggregation (bottom-up vs top-down; general government vs sub-sectors) suitable for the automatic processing of the large amount of monthly/quarterly fiscal data published nowadays by Spanish statistical authorities. Our tools are instrumental for ex-ante detection of risks to official projections, and thus can help in reducing the ex-post reputational costs of budgetary deviations. On the basis of our results, we discuss how official monitoring bodies could expand, on the one hand, their toolkit to evaluate regular adherence to targets (moving beyond a legalistic approach) and, on the other, their communication policies as regards sources of risks of (ex-ante) compliance with budgetary targets.

Keywords: Government accountability, transparency, Fiscal Forecasting.

JEL Classification: E62, E65, H6, C3, C82.

1. Introduction

Government accountability is an essential principle of democracy through which elected and non elected officials are obliged to explain to the parliament and the public in gener-

* The views expressed in this paper are the authors' and do not necessarily reflect those of the Bank of Spain or the Eurosystem. Sánchez-Fuentes acknowledges the financial support of the Spanish Ministry of Economy and Competitiveness (project ECO2012-37572) and the Regional Government of Andalusia (project SEJ 1512). Sánchez-Fuentes and Pedregal acknowledge the financial assistance of the *Instituto de Estudios Fiscales*. We thank colleagues at the Research Department of the Banco de España, and participants at the *XXI Encuentro de Economía Pública* (2014, Girona, Spain) and the *77th International Atlantic Economic Conference* (2014, Madrid, Spain), in particular Francisco Martí, Dirk Foremny and Carsten Colombier, for helpful discussions and comments. We also thank Antonio Montesinos and Mar Delgado for their help with the data. Correspondence to: Diego J. Pedregal (diego.pedregal@uclm.es); Javier J. Pérez (javierperez@bde.es); A. Jesús Sánchez-Fuentes (ajsanchezfuentes@ucm.es).

al their decisions, actions and the consequences of these decisions and actions (see e.g., IMF, 2012; Hameed, 2005; Leal *et al.*, 2008). Most developed countries do have in place a framework of political, legal, and administrative mechanisms designed to control government bodies and officials. Typically, though, these controls tend to focus on the adherence of the designed policies to the extant legal framework, and are not designed to influence the ex-ante design of policies nor the ex-post responsibility of policy-makers as regards the performance of their actions. Planning errors can have enormous influence on the economy, in particular those related to budgetary policies. For example systematic and/or sizeable budgetary forecast errors may spur waves of lack of confidence on current governmental policies, and as a consequence affect the economy as a whole via e.g. by tightening the constraints on financing channels for firms. In addition, above-fundamentals financing of the public debt imposes a burden on future generations of taxpayers. Indeed, it can be claimed that the reputational costs associated to lack of adherence of budgetary outcomes to ex-ante budgetary targets were among the group of fundamental drivers behind the recent sovereign debt crisis in the particular case of Spain, within the broader euro-area crisis.

As a consequence, a significant change in the fiscal governance framework took place since the end of 2011 in Spain, whereby an enhanced framework of national budgetary surveillance entered into force as of mid-2012, enshrined in high-ranked legal documents (Constitution and Royal Organic Law), including a huge leap forward in the availability of fiscal statistics and the procedures governing all stages of budgetary planning, including the design and the implementation phases. By now the link between the quality of fiscal frameworks and budgetary discipline is a wellproven fact from an international perspective (see, e.g. von Hagen, 2010). In the case of Spain, though, the new budgetary surveillance framework still has to prove its usefulness to control the behavior of policy makers, as regards budgetary matters, in particular in the face of upcoming electoral periods. But even assuming that policy makers were to have the will to fully implement all the legal procedures in place and were to publish timely and non-controversial real-time fiscal data, two potential weaknesses remain.

First, more information does not necessarily mean better understanding and trust from users, including private and public analysts. On the contrary, it has become increasingly difficult for private and public analysts alike to follow and interpret the continuous flow of monthly fiscal data that is currently published by official statistical agencies in Spain. Indeed, just to quote one example, in March 2012 the only monthly publication about government's budgetary execution did refer to the central government, while since the beginning of 2013 the Spanish statistical authorities do disseminate monthly information following national accounts' definitions for the central, regional and social security sectors, including individual regional governments. This is a level of dissemination of data on public accounts that has no parallel nowadays in Europe. Nevertheless, at the same time, significant and not clearly explained revisions of headline, annual past fiscal data occurred in 2012 and 2013, spurring doubts in private investors and analysts. Just to quote one example, in a Reuters' press news published as recently as November 2013, one could read: "Spain's erratic reporting of fiscal figures, especially from its regional authorities, and repeated revisions to data

have fuelled investor mistrust in the government's effort to reduce one of the euro zone's largest public deficits" ¹.

Second, the budgetary control and monitoring procedures and institutions currently in place in Spain ², pose too much weight on the ex-post adherence of policy outcomes to certain legal and administrative clauses, more than to the ex-ante design and reporting of policy actions, and the real-time monitoring of budgetary execution. This is the standard approach in continental Europe, but it remains to be seen if this approach would be able to detect in a timely fashion fiscal slippage, and to send early-warning signals that could help develop timely corrective actions.

It is against this framework that in this paper we propose a broad set of models and tools suitable for real-time monitoring of fiscal plans, including the assessment of the probability of meeting fiscal targets, that allow for a quick and efficient processing of a vast amount of incoming monthly, quarterly and annual information pertaining to most revenue and expenditure categories, and for all sub-sectors of the General Government. The models are time-series, mixed-frequencies models along the lines of Harvey and Chung (2000), Moauro and Savio (2005), Proietti and Moauro (2006), and Pedregal and Pérez (2010) ³, for which the starting point of the modeling approach is to consider a multivariate Unobserved Components Model known as the Basic Structural Model (Harvey, 1989), suitable to deal with non-stationary time series. These papers use a temporal aggregation method that relies on the information contained on related indicators observed at the desired higher frequency. The statistical treatment of structural time series models is based on the state space form and the Kalman Filter. In our case this approach allows the estimation of monthly and quarterly models using annual, quarterly and monthly observations, and permit changes over time arising from an increase in sample size. The State Space framework used allows us to build constrained forecasts in order to evaluate how feasible a given annual target might result. By setting future values of the relevant variable to the actual future policy target, we force a given model to converge to that very target, regardless of how improbable such target might be. But the interesting part of this analysis is that the model replicates a path of indicators compatible with the targets and all the information available at each moment in time. In other words, the model shows a limiting monthly/quarterly path for the indicators, necessary to meet the policy targets. We also proceed one step further and adapt the approach of Gómez and Guerrero (2006) in order to test whether official targets are compatible with the natural unconditional true forecast of a given statistical model. Such a test would produce evidence on the chances of meeting the target.

It is important to stress upfront in the paper that we see the usefulness of our models as a benchmark for the interpretation of newly available data, and not as a substitute of the in-depth analysis normally carried out by fiscal experts in policy institutions. A detailed knowledge of institutional and special factors is a key ingredient for the short-term analysis of fiscal data, which could be further exploited in conjunction with the toolkit presented in this paper ⁴. Nonetheless, being aware that budget planning and implementation is more an art than a science (as claimed, for example, by Leal *et al.*, 2008), we are at the same time convinced that looking at short-term fiscal data (i.e. data on the actual implementation of fiscal plans) through

the lens of the kind of tools and models we put forward in our paper could provide a neutral and transparent assessment of adherence of observed budgetary data against the monthly/quarterly path consistent with the achievement of annual fiscal targets⁵. On the basis of our results, we discuss how official monitoring bodies could expand, on the one hand, their toolkit to evaluate regular adherence to targets (moving beyond a legalistic approach) and, on the other, their communication policies as regards sources of risks of (ex-ante) compliance with budgetary targets, and the convenience to launch, when needed, ex-ante corrective actions.

The rest of the paper is organized as follows. In Section 2 we briefly review the related literature, stressing the contributions of the current paper. In Section 3 we describe a number of institutional issues. Then, in Section 4 we turn to the description of the data available and the publication lags of official information, to move next in Section 5 to the discussion of the methodological vagaries of our models, and present formally their potential uses for actual policy-makers. Finally, in Section 6 we show some counterfactual, empirical results to somehow justify the validity of our approach, in particular by means of truly real-time examples for the fiscal years 2011 and 2013. The former is a year in which a sizeable fiscal slippage did occur, while in the latter case the targets were broadly met. Finally, in Section 7 we provide the main conclusions of the paper.

2. Related literature

The difficulties in forecasting government borrowing needs, not only by governments, but also by other fiscal forecasters – see, e.g. Hannon (2014) – are well documented in the literature. An ex-post budgetary deviation with respect to the initial government target can be unpredictable (due to unforeseen adverse economic shocks and their impact via automatic stabilizers) or predictable if some type of strategic/electoral behaviour is considered, as the literature on politically-motivated fiscal forecasts would suggest.

In fact, official annual fiscal targets and projections have been shown in the literature to display well-documented political biases and/or large forecast errors. See for example, Strauch *et al.* (2004), Moulin and Wierds (2006), Annett (2006), Pina and Venes (2011), Jonung and Larch (2006), Leal *et al.* (2008), Merola and Pérez (2013), Frankel and Schreger (2013), and the references quoted therein. In the particular case of Spain, some papers have documented for the central government sector (*Estado*) the existence of systematic biases in budget estimates when compared to realised outcomes, for the pre-crisis period (from the 1980s till 2009). In this respect see Leal and Pérez (2009) and Pons and Solé (2001) that analyze these issues in empirical frameworks in which economic, institutional and political-economy controls are included. See also the more descriptive approach of Edo (2012). On the determinants of the budgetary deviations of regional governments see Leal and López-Laborda (2014) and Argimón and Martí (2006).

At the same time a recent strand of the literature has shown that intra-annual fiscal data, when modeled appropriately, contains extremely valuable and useful information for fore-

casting annual fiscal aggregates, enabling earlier detection of episodes of fiscal deterioration (or improvement) than traditional methods⁶. Despite this fact, monitoring public finances in the very short-run by means of high-frequency fiscal data has not been an issue traditionally tackled in the literature, even though it is usually part of the routine of practitioners⁷.

For the case of Spain, some recent contributions to the latter literature are Fernández-Caballero, Pedregal and Pérez (2012), that look at sub-national governments' budgetary data⁸, Leal and Pérez (2009; 2005), that focus on the central government sector, and Leal, Pedregal and Pérez (2010). The contribution of our paper can be seen, from the methodological point of view, as a step forward from Leal, Pedregal and Pérez (2010). Here we explore a much broader set of models, confronting their quite aggregated models with a number of single-variable models, thus enabling a clean comparison of bottom-up versus top-down approaches to the problem at hand: monitoring and forecasting general government borrowing requirements. A second contribution of our paper is that we cover the most recent crisis years. Thirdly, and more importantly, we integrate in our models the substantial amount of statistical data that has been made available by the Spanish statistical authorities since the end of 2010. Finally, and quite notably, our paper goes beyond previous only research-oriented academic contributions and provides, in addition to the research contribution, a fully-implementable toolbox usable for real-time monitoring of public finances⁹.

3. Some institutional and policy issues

Spain is currently one of the most decentralized countries in the European Union (see, e.g. Hernández de Cos and Pérez, 2013a). This is the result of a gradual transfer of responsibilities for the management of specific services from the Central Government to the Regional Governments since the beginning of the 1980s. The long process of devolution of spending and revenue capabilities to the regions has several implications for the purposes of our project. Firstly, the changes in the revenue system/spending responsibilities (substantial as of 2002 for spending, and as of 2009 for revenue) induced structural breaks in the time series of the sub-sectors of the general government, and make it difficult the evaluation of the performance of dedicated models on the basis of past data. Regional/local government data of sufficient quality and at the quarterly/monthly frequency have only been disseminated quite recently, and its usefulness in general-purpose models is still limited (see the discussion in Fernández-Caballero *et al.*, 2012). Secondly, notwithstanding the previous comment, it is possible to set up models for the sub-sectors of the general government, but more from a forward-looking point of view, i.e. it is hard to validate them on the basis of standard, out-of-sample forecasting exercises, but looking forward this does not invalidate its potential use for current policy-making.

In parallel to the process of fiscal decentralization, a number of fiscal rules were implemented in the country. In particular, the Budgetary Stability Law that entered into force in May 2012 (for details see, e.g. Hernández de Cos and Pérez, 2013b) implied a substantial step forward in two respects quite relevant for the purposes of our paper. On the one hand, the law posed quite some emphasis on ex-ante and ex-post monitoring and control proce-

dures. On the other hand, the Law has contributed fundamentally to the transparency of fiscal accounts, meaning that a huge additional amount of fiscal data has become public since May 2012. In addition to the BSL, the Spanish Parliament approved the creation of the so-called “Independent Fiscal Responsibility Authority” (AIReF), that became operational over the first half of 2014. The AIReF is in charge of issuing reports on government’s macroeconomic forecasts and budgetary plans, and monitoring adherence of the different public administrations to budgetary targets. Beyond national fiscal rules, EU countries are subject to the scrutiny of the European Commission, on the one hand, and the multilateral surveillance of the peer EU countries, following procedures defined within the confines of the so-called Stability and Growth Pact (SGP). Moreover, as a response to the 2010-2011 European sovereign debt crisis, and extensive reform of the SGP and the broader economic governance framework was adopted in 2011 (see European Commission, 2013). Six pieces of legislation (the “Six pack”) reformed both the preventive and corrective arms of the SGP.

Despite the reinforced framework of national and supra-national fiscal rules, public deficit figures deviated in many instances from ex-ante budgetary targets over the most recent 2008-2013 period. For example, deviations (higher-than-targeted public deficits) were particularly sizeable in 2011, and relatively small in 2013. We will come back to this issue when testing possible uses of our toolkit.

4. Description of the data and publication lags

4.1. The data

As mentioned above, the excellent coverage of central government budgetary execution accounts at the monthly frequency, kept on losing relevance as the devolution process to regional governments continued over the 1990s and, specially, the 2000s. The BSL posed a great deal of emphasis on the production of new budgetary execution statistics. Thus, since mid-2012 the IGAE began to regularly publish new fiscal data: quarterly accounts for all general government sub-sectors in ESA terms (since June 2012), monthly regional governments’ accounts in terms of budgetary accounts (since October 2012) and regional and Social Security monthly accounts in terms of ESA (since March 2013). Nevertheless, in almost all cases, little historical data was provided. Thus, from the point of view of the empirical analysis, most of these newly published time series are of limited use.

Nevertheless, our models incorporate all these data, given that they form the basis of real-world policy discussions. We connect them in our multi-frequency models with the rest of available data, as in many cases a wealth of historical information is available. Table 1 shows a summary of the main variables included in the dataset employed in the paper. It comprises a total amount of some 200 time series, taken from different official providers of statistics (IGAE, INE, BDSICE, Bank of Spain), and covers the period 1985-2013. The data covers the General Government sector and its subsectors. Part of the dataset is in line with national accounts’ standards, while another part follows public accounts (cash) accounting rules.

Table 1
SUMMARY OF THE DATA USED IN THE FISCAL SURVEILLANCE SYSTEM

| Non-financial series | ESA coverage* | Periodicity | Sample period | Basis | Source** | Units | | |
|----------------------------------------------|----------------------------------------------------------|----------------------|----------------------------|----------|---------------------------------------|------------|------|------------------------------------------------------|
| General government total revenue | S.13 GG | Annual, quarterly | 1985-2012 | ESA | INE (BADESPE database), IGAE | Mrd EUR | | |
| General government total expenditure | | | GG:2000Q1- 2012Q4 | | | | | |
| Central government total revenue | S.1311 CG | | | | | | | |
| Central government total expenditure | S.1311 RG | | | | | | | |
| Regional government total revenue | S.1314 SS | | Rest: 2004Q1- 2013Q3 | | | | | |
| Regional government total expend. | | | | | | | | |
| Social Security total revenue | | | | | | | | |
| Social Security total expenditure | | | | | | | | |
| Labour force survey: GG employment | | | | | | | | |
| Monthly affiliates to SS in non-market | | Monthly, | 1995- | | IGAE, INE, | | | |
| Investment of Estado | | Quarterly | 2012 | | Eurostat, other | | | |
| Public work tenders | | | | | | | | |
| AEAT public wages indicators | | | | | | | | |
| Direct taxes | | | | | | | | |
| Indirect taxes | | | | | | | | |
| VAT taxes | | | | | | | | |
| Social security contributions | | | | | | | | |
| Other revenues | | | | | | | | |
| Real government consumption | S.13 GG | | | ESA Cash | IGAE and INE | | | |
| Nominal Government consumption | | | | | | | | |
| Public investment | S.1311 subsector | | | | | | | |
| Other capital expenditure | (Estado) | | | | | | | |
| Interest payments | | | | | | | | |
| Social payments | | | | | | | | |
| Subsidies | | | | | | | | |
| Other expenditures | | | | | | | | |
| PUBLIC DEBT | | | | | | | | |
| Total cash non-financial receipts | S.1314 subsector | | | | | | | Min. of Economics /Ministry of Social Security |
| Total contributory social benefits | (Social Security System) | | | | | | | |
| Pensions | | | | | | | | Min. of Economics /Ministry of Social Security |
| Unemployment benefits | | | | | | | Cash | |
| Total expenditure | S.1314 subsector (Public State Employment Service) | | | | | | | Public State Emp. Service |
| Ceded (own) taxes of regional governments | Regional governments | Monthly | 1984-2012 | Cash | BDSICE database | | | |
| Shared taxes | | | | | | | | |
| Personal income tax | Central + | | | | | | | |
| Corporate Taxes | regional + local | Monthly | 1995-2012 | Cash | BDSICE database | | | |
| VAT | governments | | | | | | | |
| Excise duties | | | | | | | | |

(*) ESA European System Account; GG General Government; CG Central Government; LG Regional and Local Governments' aggregate; SS Social Security.

(**) INE Spanish National Institute; IGAE General Comptroller of the State Administration.

Table 2
**ASSUMPTIONS ABOUT AVAILABILITY OF DATA FOR THE COUNTERFACTUAL,
PSEUDO-REAL-TIME, FORECASTING EXERCISES**

| | First quarter (March) | Second quarter (June) | Third quarter (September) | Fourth quarter (December) |
|-------------------------|--------------------------|-----------------------------------|------------------------------|------------------------------|
| Annual data up to | Year $t - 2$ | Year $t - 1$ (April) | Year $t - 1$ | Year $t - 1$ |
| Quarterly data up to | Third quarter $t - 1$ | Fourth quarter $t - 1$ (April) | First quarter t | Second quarter t |
| Monthly data up to | January t | April t | July t | October t |

4.2. Publication lags and timing convention

Annual fiscal outturns for a given year t are published at the very end of March of year $t + 1$. Quarterly non-financial accounts for the General government and all its sub-sectors are published regularly with a delay of 90 days. Monthly data for the State sector (“Estado”) are published with a lag of one month. Also with a lag of one month are published the data on shared (by “Estado”, regional and local governments) taxes’ collection, and social security system outturns, in both cases in cash terms. As regards the newly available information, monthly national accounts data for the Central Government, the regions and the Social Security, are published with a delay of broadly two months.

Table 3
**FORECASTING THE 2011 FISCAL YEAR IN REAL TIME: INFORMATION
AVAILABLE IN FEBRUARY 2011, JUNE 2011 SEPTEMBER 2011, NOVEMBER 2011
AND FEBRUARY 2012**

| | February 2011 | June 2011 | September 2011 | November 2011 | February 2012 |
|-------------------------|-----------------------|------------------------|-----------------------|------------------------|------------------------|
| Annual data up to | Year 2009 | Year 2010 | Year 2010 | Year 2010 | Year 2010 |
| Quarterly data up to | Third quarter 2010 | Fourth quarter 2010 | First quarter 2011 | Second quarter 2011 | Second quarter 2011 |
| Monthly data up to | December 2010 | March 2011 | July 2011 | September 2011 | November 2011 |

Given the absence of comprehensive historical data records of real-time data, it is not possible to fully re-create the real-time nominal fiscal series that would have been available at each point in time. Thus, our baseline forecasting illustration will be based on the data series available in October 2013. In this respect, the empirical exercises would have some counter-factual features in that data revisions might have affected the lessons drawn from the application of the toolbox in relation to a today’s re-creation. Nonetheless, in order to replicate the real-time constraints faced by real-time fiscal policy analysts, we adopt the timing rules displayed in Table 2, following the standard dates of dissemination of data at the dif-

ferent frequencies. We deem this convention as a fair heuristic representation of reality, on average. Nonetheless, in addition to the pseudo-real-time exercises, we were able to assemble two truly real-time datasets to illustrate the real-time usefulness of our toolbox in two particular episodes: the fiscal years 2011 and 2013. In Table 3 we show the exact information available at each specific date of reference for the 2011 exercise, being similar the information set for the 2013 example.

5. Main features of the toolkit

5.1. A suite of models

Given the different sampling frequencies of the time series included in our dataset, we estimate multivariate, mixed-frequencies models, of the unobserved components type, following the lead of Harvey and Chun (2000). The particular unobserved components model considered is known as the Basic Structural Model (Harvey, 1989). A given time series is decomposed into unobserved components which are meaningful from an economic point of view: trend, T_t , seasonal, S_t , and irregular, e_t . In this regard, the models are flexible enough to model non-stationary time series, the standard case when dealing with economic time series. All the details of the modeling approach are described in Appendix A. This type of models have been used with success in the field of fiscal forecasting (Onorante *et al.*, 2009; Paredes *et al.*, 2014; Pedregal and Pérez, 2010; Leal *et al.*, 2010; Fernández-Caballero *et al.*, 2012). Their flexibility allows us to accommodate a number of policy-relevant exercises. We develop a number of models that look at the data from different, complementary angles. In particular:

- Model 1 (“Joint model”): this is an aggregated model for total revenues and expenditures of the general government sector in national accounts. Three “indicators” (public accounts series pertaining to the central government and the social security system) are used for each general government aggregate, but the model is estimated jointly so that the dynamics of revenues and expenditures are jointly determined. This is in line with the literature, that predicts a tight linkage between revenues and expenditures (in the form of bi-directional causality relationships). The model is set up and estimated at the monthly frequency.
- Model 2 (“Direct model”): this is an aggregated model for the general government sector balance in national accounts terms. Short-term indicators are the monthly balances of the sub-sectors of the general government, in public accounts’ terms. The model is set up and estimated at the monthly frequency.
- Model 3 (“Disaggregated model, components of total revenues and total expenditures”): single models for each one of the sub-components of total general government revenues and expenditures are produced independently, and forecasts for the general government borrowing requirement are produced by bottom-up aggregation. The models are set up and estimated at the quarterly frequency.

- Model 4 (“Disaggregated model, total revenues and total expenditures”): single models for total general government revenues and expenditures are produced independently, and forecasts for the general government borrowing requirement are produced by subtraction. The models are set up and estimated at the quarterly frequency.
- Model 5 (“Joint, sectoral models”): models by each one of the sub-sectors of the General Government (Central, Regions, Social Security) along the lines of Model 1. Forecasts for the general government borrowing requirements are produced by aggregation.
- Model 6 (“Direct, sectoral models”): models by each one of the sub-sectors of the general government (Central, Regions, Social Security) along the lines of Model 2. Forecasts for the general government borrowing requirement are produced by aggregation.

The classes of models considered allow us in particular to address the questions of interest, while at the same time providing information on the “bottom-up versus top-down” dichotomy and the “aggregate general government approach versus approach by sub-sectors” dichotomy. All the models can be written in a common general form as described in Appendix A. Our approach is closely related to Harvey and Chung (2000), Moauro and Savio (2005), and Proietti and Moauro (2006). These papers use a temporal aggregation method that relies on the information contained on related indicators observed at the desired higher frequency. The statistical treatment of structural time series models is based on the state space form and the Kalman Filter (see Harvey, 1989). In our case this approach allows the estimation of monthly/quarterly models using annual, quarterly and monthly observations, and permit changes over time arising from an increase in sample size.

5.2. Evaluating the probability of meeting the budgetary target

The methodology used, i.e. the State Space framework, allows for some easy and relevant exercises. It is straightforward to build constrained forecasts in order to evaluate how feasible a given target might result. By setting future values of the indicated variables to the actual future targets of a given public agency, we force the model to converge to that target (due to the built-in time constraints in the model), regardless of how absurd or improbable such targets might be. But the interesting part of this analysis is that the model replicates a path of indicators compatible with the targets and all the information available at each moment in time. In other words, the model shows a limiting monthly path for the indicators necessary to meet the future annual targets issued by the agency.

Within this framework, we can proceed in a more formal way to test whether a the government target (or a given forecast provided by any public or private agency) is compatible with the natural unconditional true forecasts of a selected model. Such a test would produce evidence in favor or against the chances of meeting that very target. In order to implement such a test, we adapt the approach in Gómez and Guerrero (2006) for a general State-Space

model. Let's assume all noises in the model are Gaussian, R is a vector of m targets and Y_F are the forecasts of the model, with covariance matrix Σ_y . Meeting the constraints imply that $R = Y_F + u$, with $u \sim N(0, \Sigma_u)$, a noise that is assumed independent of any other noises in the model. Then, the formal compatibility test consists of evaluating the distance $d = R - Y_F$, which distribution is $N(0, \Sigma_y + \Sigma_u)$. For the estimated values in a particular model the compatibility test is based on the statistic

$$K = d' (\Sigma_u + \Sigma_y)^{-1} d / m \sim F_{m,r} \quad (1)$$

where r is the length of the series minus the number of parameters involved in the model. Gómez and Guerrero (2006) distinguish two cases: (i) the targets are considered as binding constraints, i.e. the constraints ought to be met exactly (this is equivalent to saying that $u = 0$); (ii) the targets are unbinding constraints, i.e. the constraints are met "statistically", not exactly with $u \neq 0$, and a feasible estimation of Σ_u is necessary in order to carry out the test. In our paper we are only interested in case (i). Indeed, it turns out that in the present context, where a set of given forecasts are compared to targets issued by a government that is committed to achieve them, such targets ought to be taken as binding constraints, mainly because the target is achieved only when the true value of the variable exceeds that of the target. Thus, in our case we use the distribution of the statistic above for the case in which $\Sigma_u = 0$, to calculate probabilities of meeting the targets.

In the latter regard, a distinction between public deficit and revenues on the one hand, and public expenditures on the other, has to be taken into account. Certainly, meeting a deficit or revenue target means in practice achieving a value greater or equal to the target, as a matter of fact, the further is the unconditional forecast above the target implies a greater probability of meeting such a target. Formally, the probability of meeting the target would translate into $P(Y_F \geq R)$. If the unconditional forecast hits exactly the target, then the probability of meeting the target is 0.5. On the other hand, for expenditures further unconditional forecasts below the target is evidence of a high probability of meeting the target, formally we ought to calculate $P(Y_F \leq R)$.

6. Empirical illustrations

6.1. Some general remarks

We perform a rolling forecasting exercise in which the selection of the forecast origin and the information set available at each date are carefully controlled for. In particular we evaluate the forecasts generated from four forecast origins per year from March 1999 to December 2012, which makes up to 14×4 projections at each forecast horizon. The first forecast origin is March 1999, and following the timing convention outlined before (see Table 2) the annual information available covers up to the year 1997, the quarterly information up to 1998:Q3, and the monthly information up to January 1999. The second forecast origin is June 1999, with annual information up to 1998, quarterly up to 1998:Q4 and monthly up to January-April 1999.

Then we move the forecast origin to September 1999 and so on and so forth until December 1999. We focus on the forecast performance for annual projections, i.e. forecasts generated from each forecast origin for the end of the current year, as this is the main horizon of use for mechanical, time-series based forecast. From the point of view of a practitioner, forecasts of fiscal variables for a horizon longer than the current year is of less importance. Our tools are developed to monitor the budget, and the latter, in the case of Spain, follows an annual cycle.

For the nominal public balance, the forecast error committed for year t by model J from forecast origin Q is defined as

$$\varepsilon_{t,Q}^{BAL,J} \equiv (BAL_t - BAL_{t-1}) - (E_{\Omega,Q}^J[BAL_t] - E_{\Omega,Q}^J[BAL_{t-1}]) \quad (2)$$

where Ω makes reference to the information set available at the time of generating a given forecast, as described in Table 2. For revenue and expenditure items, the error committed in year t for item I by model J from forecast origin Q is defined as

$$\varepsilon_{t,Q}^{I,J} \equiv \left(\frac{I_t}{I_{t-1}} - 1 \right) * 100 - \left(\frac{E_{\Omega,Q}^J[I_t]}{E_{\Omega,Q}^J[I_{t-1}]} - 1 \right) * 100 \quad (3)$$

We compute two standard quantitative measures of forecasting performance for a number of pseudo-real-time forecasting exercises. On the one hand, the ratio of the Root Mean Squared Errors (RMSE) of the different alternative models with respect to an annual random walk (i.e. no-change) alternative. On the other hand, we also look at a qualitative measure of forecast performance, namely, whether the predicted change coincided or not with the actual change observed in the variable of interest. We also present, as discussed in a previous Section, a truly real-time exercise, focused on the 2011 fiscal year, given the relevance of the budgetary deviation observed that year.

6.2. Bottom-up vs top-down models

The results of a first exercise are presented in Table 4. In that table we show the relative root mean squared error of our models compared to the annual random walk extrapolation for a number of cases: (i) aggregate of the forecast errors generated for the whole year from all forecasts origins (baseline); (ii) forecasts errors computed on the basis of forecasts computed taking as forecasts origin the first quarter (Q1), the second (Q2), the third (Q3) and the fourth (Q4); (iii) these exercises are presented for the whole sample used for the rolling forecasting exercise (“Full sample”, 1999-2012), and for the crisis sample (2008-2012). The following messages can be highlighted from Table 4 for a subset of the results obtained, namely the aggregated results for the general government balance, revenues and expenditures.

First, when looking at the full sample, and pooling all forecast errors (resulting from forecast origins Q1 to Q4), the most aggregated models (i.e. those that model directly the budget deficit), models 1 (general government) and model 6 (sub-sectors), are the best. All

other models are close to these ones, with the exception of model 4. This relative ranking of models is broadly kept when looking into forecasts from each origin (Q1 to Q4, taken individually). Thus, with the exception of models 3 versus 4, in the other two cases more aggregate models outperform more disaggregated models. This can be taken as evidence that bottom-up approaches are not necessarily better than top-down ones, at least as regards forecast accuracy. Of course, in real-time, bottom-up approaches provide the advantage of giving a more comprehensive view, which can be an asset in cases like the current one in which overall performance across models is not overwhelmingly different. The main results on bottom-up versus top-down holds when looking into sub-samples.

Table 4
QUANTITATIVE FORECASTING PERFORMANCE OF ALTERNATIVE MODELS:
SUMMARY OF RESULTS FOR THE GENERAL GOVERNMENT BALANCE,
REVENUE AND EXPENDITURE*

| | General Government balance | | | | | | Total revenues | | | | Total expenditures | | | |
|-------------|----------------------------|------|------|------|------|------|----------------|------|------|------|--------------------|------|------|------|
| | M1 | M2 | M3 | M4 | M5 | M6 | M1 | M3 | M4 | M5 | M1 | M3 | M4 | M5 |
| Full Sample | 0.88 | 0.91 | 0.96 | 1.19 | 1.01 | 0.90 | 1.04 | 1.01 | 1.15 | 1.00 | 0.95 | 0.93 | 0.86 | 1.31 |
| Q1 | 0.95 | 0.99 | 1.05 | 1.20 | 1.18 | 0.99 | 1.15 | 1.18 | 1.17 | 1.03 | 0.98 | 0.89 | 0.78 | 1.23 |
| Q2 | 1.12 | 1.17 | 1.16 | 1.43 | 1.21 | 1.09 | 1.38 | 1.10 | 1.33 | 1.14 | 1.17 | 1.28 | 1.04 | 1.68 |
| Q3 | 0.70 | 0.75 | 0.84 | 1.19 | 0.77 | 0.73 | 0.87 | 0.85 | 1.15 | 0.96 | 0.80 | 0.80 | 0.91 | 1.21 |
| Q4 | 0.55 | 0.47 | 0.62 | 0.88 | 0.58 | 0.66 | 0.50 | 0.71 | 0.92 | 0.84 | 0.74 | 0.72 | 0.78 | 1.15 |
| Crisis | 0.84 | 0.87 | 0.95 | 1.20 | 1.00 | 0.90 | 0.87 | 1.00 | 1.15 | 0.96 | 0.67 | 0.56 | 0.86 | 1.25 |
| Q1 | 0.91 | 0.97 | 1.05 | 1.20 | 1.18 | 0.99 | 1.04 | 1.18 | 1.16 | 1.03 | 0.61 | 0.61 | 0.76 | 1.11 |
| Q2 | 1.10 | 1.11 | 1.13 | 1.43 | 1.19 | 1.09 | 1.05 | 1.11 | 1.34 | 1.11 | 0.96 | 0.68 | 1.13 | 1.70 |
| Q3 | 0.68 | 0.72 | 0.84 | 1.20 | 0.75 | 0.72 | 0.72 | 0.82 | 1.16 | 0.93 | 0.68 | 0.45 | 0.94 | 1.23 |
| Q4 | 0.48 | 0.42 | 0.62 | 0.89 | 0.55 | 0.63 | 0.42 | 0.66 | 0.91 | 0.68 | 0.40 | 0.29 | 0.70 | 1.18 |

(*) The numbers in the table are the ratios of Root Mean Squared Errors of the errors obtained with each model (labeled models 1 to 6) with respect to an annual random walk approach (no-change baseline). Forecast errors are computed as follows. For the nominal balance, the error committed for year t by model J from forecast origin Q is defined as $\varepsilon_{t,Q}^{BAL,J} \equiv (BAL_t - BAL_{t-1}) - (E_{\Omega,Q}^J[BAL_t] - E_{\Omega,Q}^J[BAL_{t-1}])$. For revenue and expenditure items, the

error committed in year t for item I by model J from forecast origin Q is defined as $\varepsilon_{t,Q}^{I,J} \equiv \left(\frac{I_t}{I_{t-1}} - 1 \right) * 100 - \left(\frac{E_{\Omega,Q}^J[I_t]}{E_{\Omega,Q}^J[I_{t-1}]} - 1 \right) * 100$. Ω makes reference to the information set available at the time of generating a given forecast, as described in Table 2.

Second, in general, the forecast accuracy of all models is better in the crisis sample than in the “expansion” one. This may reflect the fact that the models can do a fair job in periods of significant changes, while in a period with no fiscal stress and persistent economic growth, it is more difficult to beat a simple extrapolation of the past.

Third, across quarters, the forecasting performance of all models improves when more information about revenue collection and the implementation of spending plans kicks-in. This is quite

clear in the second half of the year compared to forecasts prepared in the first half. In particular, in Q3 a fair amount of information for the first half of the year is assumed to be available, but only the first quarter of the general government accounts, while in Q4 the first half of the year is fully known. For projections prepared from forecast origin Q2 things are quite different. In our timing convention this is the quarter in which the annual figure of year $t - 1$ is known. This seems to create a discontinuity in how models process incoming information, as forecast accuracy is worse than Q1-based forecasts, a fact that may be linked to the realization of past data revisions, including the appearance of “hidden spending” not reflected in monthly/quarterly indicators.

Fourth, when looking at revenue and expenditure errors, the same general results as regards full sample versus crisis-sample forecasts, and as regards first semester versus second semester forecasts, hold. Interestingly, models add more information compared to the simple random walk baseline in the case of expenditures than in the case of revenues, i.e. in general relative RMSEs tend to be lower across models, though not in all cases. This provides some evidence on the ability of models to accommodate purely within-the-year discretionary policy changes.

Some interesting results can be highlighted from the particular set of revenue/expenditure projections. In the case of total revenues: (i) the aggregated, joint revenue-expenditures model (model 1, “M1”) tends to be the best performer, in particular in the case of the crisis period; this is the case versus the bottom-up approach, M3, and model M4 that does not exploit the link between revenues and expenditures; (ii) as regards the other models, M3 seems to be the best in general, even though it is not clear-cut along all the dimensions considered (full-sample/pre-crisis, across quarters). As regards total expenditures: (i) bottom-up approaches seem to provide (marginal) better results, being clear from the comparison of M3 versus M1, which makes sense given that, in particular, fiscal consolidations tend to have a differentiated profile for the different spending components; while social payments tend to increase in crisis times (unemployment benefits) or to stay, at most, stable (pensions), other components like public investment, or the wage bill, tend to move in the opposite direction in fiscal adjustment periods this is a differentiated element when compared to government revenues, that are subject to similar macroeconomic shocks, even though tax hikes can be uneven across revenue aggregates; (ii) the aggregation of sectoral models (M5) performs quite badly in the case of TOE, which is not surprising given the fact that central government transfers to the rest of the sectors, in particular to the social security, tend to occur during the year, distorting, thus, the genuine signals of sectoral data.

Beyond the comparison of alternative models across several dimensions of this subsection, it is by now a proven fact in the literature that the combination of alternative models tends to outperform individual models. In the next subsection of the paper we exploit that dimension of our models.

6.3. The usefulness of the combination of models

In Tables 5 and 6 we compare the performance of the combination of models’ forecasts with the forecasts of the European Commission (EC henceforth). As shown in Artis and

Marcellino (2001) and Keereman (1999), the forecast record of the EC is among the best of the international organizations producing regular forecasts for European countries, and in particular Spain (others include the International Monetary Fund and the OECD). EC forecasts tend to make use of all the information available at the time the forecasts are done, not only observed data, but also all available, forward-looking information on budgetary plans, including additional corrective packages enacted by the governments in the course of the year. EC forecasts are based on a bottom-up approach. In addition, EC fiscal forecasts use both macroeconomic models and expert judgement. That is why checking the performance of the models (specifically, a combination of them) against EC forecasts should be quite a demanding criterion. Even bearing in mind that we are comparing against a difficult-to-beat benchmark, our objective with this exercise is to check the usefulness of the models to complement an approach that takes into account backward and forward-looking information alike, as in the case of the EC forecast.

Table 5
QUANTITATIVE FORECASTING PERFORMANCE^a: COMBINATION
OF MODELS' FORECASTS^b (MEAN AND MEDIAN) VERSUS
EUROPEAN COMMISSION FORECASTS^c

| | Government balance | | | Government revenues | | | Government expenditures | | |
|--------------------|--------------------|--------|------|---------------------|--------|------|-------------------------|--------|------|
| | Mean | Median | EC | Mean | Median | EC | Mean | Median | EC |
| Full Sample | 0.87 | 0.91 | 0.82 | 0.90 | 0.88 | 1.47 | 0.85 | 0.81 | 0.68 |
| Q1 | 0.98 | 1.04 | 0.85 | 1.01 | 1.00 | 1.13 | 0.86 | 0.75 | 0.51 |
| Q2 | 1.06 | 1.07 | 0.88 | 1.03 | 1.04 | 1.65 | 1.06 | 1.00 | 0.82 |
| Q3 | 0.74 | 0.80 | 0.88 | 0.83 | 0.81 | 1.65 | 0.73 | 0.82 | 0.82 |
| Q4 | 0.49 | 0.49 | 0.64 | 0.58 | 0.51 | 1.55 | 0.67 | 0.71 | 0.72 |
| Crisis | 0.87 | 0.91 | 0.81 | 0.87 | 0.86 | 1.49 | 0.72 | 0.68 | 0.59 |
| Q1 | 0.98 | 1.04 | 0.84 | 1.01 | 1.00 | 1.13 | 0.70 | 0.64 | 0.44 |
| Q2 | 1.05 | 1.06 | 0.86 | 0.99 | 1.01 | 1.68 | 0.96 | 0.89 | 0.73 |
| Q3 | 0.74 | 0.79 | 0.86 | 0.81 | 0.78 | 1.68 | 0.67 | 0.70 | 0.73 |
| Q4 | 0.49 | 0.48 | 0.64 | 0.51 | 0.44 | 1.56 | 0.51 | 0.46 | 0.66 |

^a The numbers in the table are the ratios of Root Mean Squared Errors of the errors obtained with each alternative with respect to an annual random walk approach (no-change baseline).

^b For the nominal balance, the error committed for year t by model J from forecast origin Q is defined as $\varepsilon_{t,Q}^{BAL,J} \equiv (BAL_t - BAL_{t-1}) - (E_{\Omega,Q}^J[BAL_t] - E_{\Omega,Q}^J[BAL_{t-1}])$. For revenue and expenditure items, the

error committed in year t for item I by model J from forecast origin Q is defined as $\varepsilon_{t,Q}^{I,J} \equiv \left(\frac{I_t}{I_{t-1}} - 1 \right) * 100 - \left(\frac{E_{\Omega,Q}^J[I_t]}{E_{\Omega,Q}^J[I_{t-1}]} - 1 \right) * 100$. Ω makes reference to the information set available at the time of generating a given forecast, as described in Table 2.

^c European Commission forecasts are taken from AMECO real-time vintages. For the nominal balance, the error committed for year t from forecast origin Q is defined as $\varepsilon_{t,Q}^{BAL,EC} \equiv (BAL_t - BAL_{t-1}) - (E_Q^{EC}[BAL_t] - E_Q^{EC}[BAL_{t-1}])$.

For revenue and expenditure items, the error committed in year t for item I by model J from forecast origin Q is defined as $\varepsilon_{t,Q}^{I,EC} \equiv \left(\frac{I_t}{I_{t-1}} - 1 \right) * 100 - \left(\frac{E_Q^{EC}[I_t]}{E_Q^{EC}[I_{t-1}]} - 1 \right) * 100$

In Table 5 we show the relative RMSE of each alternative with respect to the random walk approach. Overall, for the full sample and when all the errors from all forecast horizons are pooled, EC government balance forecast errors are lower than the mean and the median of the alternative models. This is also the case as regards full sample government expenditure errors, while in the case of public revenue the opposite happens. This full-sample picture also holds when looking at the crisis sample. Quite interestingly, though, Q3 and Q4-based forecast errors are systematically lower for the combination of models versus EC forecast, both for the whole sample and the crisis sample. This means that as soon as a sufficient amount of data is available on the implementation of spending plans and/or the behavior of revenue collection (under our quite restrictive timing convention in Q3 only the first quarter of the headline general government variable, and half-year in the case of indicators are available) the models are able to process and extrapolate these data in a quite informative way. By this we mean, in a way that even tends to outperform a full information approach that incorporates in an explicit manner forward-looking elements (policy measures affecting future quarters). In particular, in Q4 the mean and the median of models displays a remarkable forecast accuracy. Turning to government revenues, it is surprising the relative bad forecast accuracy of EC forecasts for government revenue growth projections, a result that is dominated by the significant forecast errors around GDP turning points, related to the double-dip crisis. Turning now to Table 6, we present the percentage of correctly predicted changes in the case of government balance, and the percentage of correctly predicted signs of the growth rate in the case of government revenues and expenditures. The qualitative results presented there display similar messages to the ones discussed before on quantitative measures.

6.4. More on the properties of models' forecast errors

We provide now some additional results to better understand some of the conditions under which the toolkit provides more accurate/reliable signals, according to the previous counterfactual, pseudo real-time exercises. This complements the previous analysis in which we have looked at quantitative and qualitative forecast performance measures by forecast origin and by crisis/non-crisis periods. Thus, in Table 7 we pool all models' forecast errors and compute the resulting probability distribution function, distinguishing between government balance, revenue and expenditure. What is more, we zoom in each case the probability distribution function by forecast origin (Panel A) and good times versus bad times (Panel B).

Some results are worth highlighting. First, most of the probability mass is concentrated in the bins from -2% to $+2\%$ of GDP. Indeed, when looking at the complete pool, the numbers are 64%, 71% and 88%, for balance, revenue and expenditure errors, respectively. These percentages increase substantially (i.e. the occurrence of high deviations decrease) when considering Q4-based forecast errors (77%, 83% and 92%) and "boom periods" (87%, 92% and 95%).

Table 6
QUALITATIVE FORECASTING PERFORMANCE^a: COMBINATION
OF MODELS' FORECASTS^b (MEAN AND MEDIAN) VERSUS
EUROPEAN COMMISSION FORECASTS^b

| | Government balance | | | Government revenues | | | Government expenditures | | |
|--------------------|--------------------|--------|------|---------------------|--------|------|-------------------------|--------|------|
| | Mean | Median | EC | Mean | Median | EC | Mean | Median | EC |
| Full Sample | 45.8 | 43.8 | 58.3 | 68.8 | 68.8 | 39.6 | 41.7 | 47.9 | 62.5 |
| Q1 | 25.0 | 25.0 | 58.3 | 50.0 | 50.0 | 50.0 | 41.7 | 41.7 | 75.0 |
| Q2 | 33.3 | 33.3 | 58.3 | 66.7 | 75.0 | 33.3 | 16.7 | 50.0 | 50.0 |
| Q3 | 58.3 | 41.7 | 58.3 | 83.3 | 75.0 | 33.3 | 58.3 | 50.0 | 50.0 |
| Q4 | 66.7 | 75.0 | 58.3 | 75.0 | 75.0 | 41.7 | 50.0 | 50.0 | 75.0 |
| Crisis | 50.0 | 40.0 | 35.0 | 75.0 | 75.0 | 40.0 | 50.0 | 55.0 | 70.0 |
| Q1 | 20.0 | 20.0 | 20.0 | 40.0 | 40.0 | 40.0 | 40.0 | 40.0 | 80.0 |
| Q2 | 40.0 | 40.0 | 40.0 | 80.0 | 60.0 | 40.0 | 40.0 | 60.0 | 60.0 |
| Q3 | 60.0 | 40.0 | 40.0 | 80.0 | 100.0 | 40.0 | 60.0 | 60.0 | 60.0 |
| Q4 | 80.0 | 60.0 | 40.0 | 100.0 | 100.0 | 40.0 | 60.0 | 60.0 | 80.0 |

^a The numbers in the table are the percentage of correctly predicted changes in the case of the government balance, and the percentage of correctly predicted signs of the growth rate in the case of government revenues and expenditures.

^b See footnotes to Table 5.

Second, the probability distributions are not symmetric. When looking at the whole pool for government balances, there is more probability mass on the negative side of the distribution (some 55%), i.e. public deficits tend to be underestimated. This is due to a higher probability of negative revenue errors (60% of the mass), i.e. optimistic forecasts of revenues, and despite the fact that government expenditures also tend to be overestimated (negative errors represent 55% of the mass), i.e. do present a conservative bias. The within-the-year evolution of this asymmetry is interesting: despite the forecast origin being Q1, Q2-Q3 or Q4, the amount of negative errors for the government balance (optimistic forecast) outweighs that of positive errors (on average 54% vs. 46%). As regards revenue forecasts, these tend to be optimistic for all forecast origins, even though the percent declines as the end of the year approaches (from 58% for Q1-based forecast, to 55% for Q4-based). Regarding expenditure forecasts, the conservative bias shows up as of Q2 forecasts. Splitting the sample between boom and crisis periods is also interesting from the “asymmetry” point of view: the overall result that negative government balance errors predominate (i.e. optimistic deficit bias) is dominated by the boom period (59% of occurrences are negative) while over the crisis models’ forecasts errors were more evenly distributed (49% negative, 51% positive). Along the same lines, the observed overall overestimation of revenues is dominated by boom years (65% of positive errors), but it is also present, to a lesser extent, in crisis years (54%). In line also with the observed result for the budget balance, expenditure forecast errors follow quite different patterns in boom and crisis years: in the former, positive errors (underestimation of expenditures) dominate (56% of the total), while over the crisis expenditures were overestimated in 75% of the cases (negative errors, conservative bias). The latter may reflect the successive implementation of expenditure adjustment plans over the crisis.

Finally, it is worth mentioning that during the crisis the uncertainty surrounding model-based forecasts increases substantially. This is to be expected and is in line with results reported for general macroeconomic forecasts as well as for non-model-based budgetary projections, i.e. private and public sector analysts' ones. In our case this is quite evident from "gross tails" in the distribution of crisis-years' government balance forecast errors. The result for the balance is dominated by the profile of government revenue errors. On the contrary, government expenditure trends, as captured by the models, did not display that huge dispersion: in fact, in crisis years 96% of the expenditure forecast errors were within the -4% to $+4\%$ interval, compared to only 64% in the case of revenues. In the "boom" years the percentages turned out to be 100% and 98% for expenditures and revenues, respectively.

6.5. Real-time forecasting exercises

In this subsection we present two additional exercises, truly real-time ones, i.e. based on the exact dataset available at each point in time. We select two specific periods: the fiscal years 2011 and 2013. The results are shown in Figures 2 and 3.

We firstly focus on the fiscal year 2011, a very specific year given the sizeable forecast deviation committed by all public and private sector analysts. As shown in Figure 1, the government managed to anchor the expectations of private forecasters (summarized by the mean of the "Consensus Forecast" panel) and independent public analysts (European Commission, OECD and IMF) around the target of 6% of GDP for the whole year. Indeed, the anticipated budgetary deviation was on average less than 1% of GDP, even for the most extreme forecaster. As a consequence, all forecasters committed substantial forecast errors, as the deficit figure initially published in Spring 2012 for the year 2011 was 8.5% of GDP.

As discussed, from each forecast origin, and conditional on the short-term information available, we can compute unconditional forecasts and also the consistency of government targets with these forecasts. In particular, in Figure 2 we present in each single box how the unconditional forecasts (in white color) are updated with additional information, in relation to the actual final data (dots), the targets (thick line red) and the uncertainty in each case (fan-chart up to 99% confidence). The two top rows of charts of Figure 2 show the results obtained with models 1 and 2 (for the public deficit), and the third and fourth rows show total revenues and total expenditures with model 1. Numbers in parentheses in each title are probabilities of meeting the target, i.e. $P(Y_F \geq R)$ for deficits and revenues, and $P(Y_F \leq R)$ for expenditures, as was stated in the previous section.

The two left-upper panels show government balance projections for 2011 from forecast origin in February 2011. The information available was quite scarce at that point in time: the annual figure for 2009, quarterly information up to the third quarter of 2010 and monthly indicators for the month of December in some cases. Thus, it is not surprising to see that confidence bands generated with both models were quite broad, reflecting a extremely high uncertainty surrounding the projections (white lines), and a central scenario of further

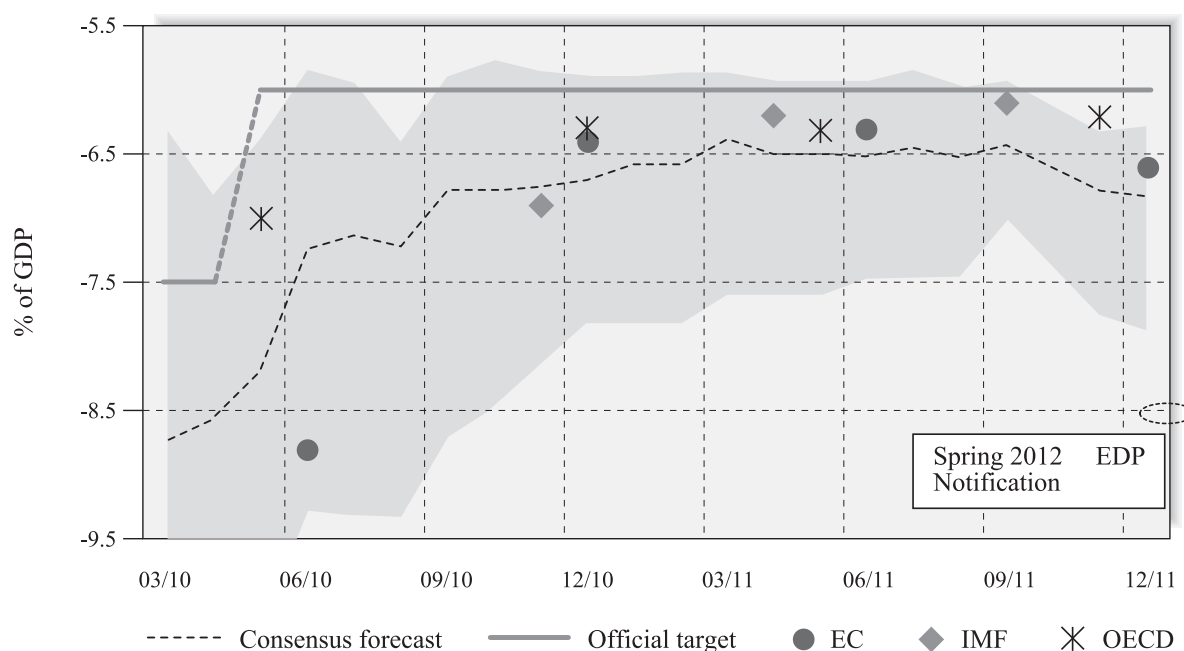


Figure 1. The 2011 fiscal targets (red, horizontal lines): forecast prepared by different private and public institutions for the general government balance in 2011, from different forecast origins

Source: Stability and Convergence Programmes, Consensus Forecast, OECD, IMF and EC Spring and Autumn reports, several years. The grey area represents the distance between the max and min forecast within the Consensus Forecast pool of forecasters.

deterioration of the budget balance. In June 2011, the information available increased substantially: the first estimate for the year 2010 was published and short-term indicators covered up to March 2011. Model 1 (first line of charts) indicated at that moment an improvement in the budget balance in 2011 compared to 2010, but still the probability associated with meeting the target was 24% percent, while in the case of Model 2 (joint model of revenues and expenditures) the probability assigned to meeting the target was zero, i.e. the target was out of the confidence bands of the model forecast. In the latter case, the probability assigned to meeting the government revenue target, though, was not null, but was as small as 5%. The publication of the general government figures for Q1 and of short-term indicators up to July (September 2011 forecast origin) improved marginally the revenue projection, while in the case of the direct-deficit-model (model 1) the probability assigned to meeting the target was around 10%, but still, in qualitative terms the forecast signalled an improvement in the government balance, a direction consistent with the objective of the government.

In November 2011 (information set: second quarter of general government variables, short-term indicators up to September), though, the situation change considerably. Both models were assigning a zero probability to the event “the target is met”, and at the same time, both models signalled to basically no improvement in the budget balance compared to the one observed in 2010. In the case of model 2, estimated government revenues for 2011 were for a significant

drop compared to the previous year, and the balance was estimated at more or less the same nominal amount as in 2010 due to the projected expenditure restraint, as in the official plans (red thick line). It is worth mentioning that in the last quarter of 2011 the double-dip in real GDP growth was starting to be evident, a fact that was visible in government revenue aggregates. Compared to the actual realization of deficits, revenues and expenditures, the estimates computed by those models in real-time was indeed quite accurate. This is even more clear when looking at the forecast labeled “February 2012”, that incorporated data for general government, headline variables up to Q2 2011, and short-term indicators up to November 2011.

The real-time recreation of the use of two of our suggested models for the 2011 episode is quite illustrative about the potential uses of our system. The bold signals of the models were quite clear, in particular after the summer. Since November, the two selected models were signalling clearly to a around 2% of GDP deviation with respect to the official, government target. In the case of the revenue/spending model, in addition, the slippage was almost fully related to revenue shortages, while expenditures, even being forecast to be above the target, were relatively close in quantitative terms. These results contrast with the “herding behavior” observed in Figure 1, and indicates that the data on the implementation of revenue targets were already hinting towards a sizeable slippage, despite corrective measures adopted at the end of the summer. Indeed, in August a broadening of corporate tax bases was approved, and entered into force in Q4, but as not able to compensate the strong, downward trend in tax collection that was visible at least since November 2011 (with data available up to September at the maximum).

We perform a second real-time illustration for the fiscal year 2013. The results are shown in Figure 3. Contrary to the discussion above, the year 2013 is one in which public finances followed a more stable path during the year, and the target was broadly met. It is worth mentioning that in mid-2012 the Spanish Government introduced a significant fiscal consolidation package that drastically changed the course of policies. Data for the second half of 2012 showed a significant improvement in the public deficit, in particular as regards the expenditure side. The first column of charts seems to present an overreaction by the models, with the information available in March, i.e. general government data up to 2012Q3 and basically no information regarding 2013. In particular, the models were anticipating an undershooting of government expenditures compared to the target, while revenues were still projected to be well below the official forecast. The evolution of forecasts as information pertaining to 2013 becomes available (i.e. as we move to the right in the figure) is quite revealing. As regards government expenditures, model forecasts showed limited variability and centered around the target since mid-year, a target that was met at the end. Regarding revenues, the sequential entry into force of tax increases and the mild turn in the business cycle implied that model forecasts were adjusting upwards, and newly available incoming information confirmed that the implementation of the plans was leading to some success in raising revenues. Looking at public deficits, and in line with the previous commentary of revenues and expenditures, the models were signalling already since June 2013 (second column of charts, i.e. when the 2012 annual figure was known, as well as budgetary indicators for the first quarter of 2013) limited risks for budgetary deviations, a signal that was kept consistently as new information was taken on board.

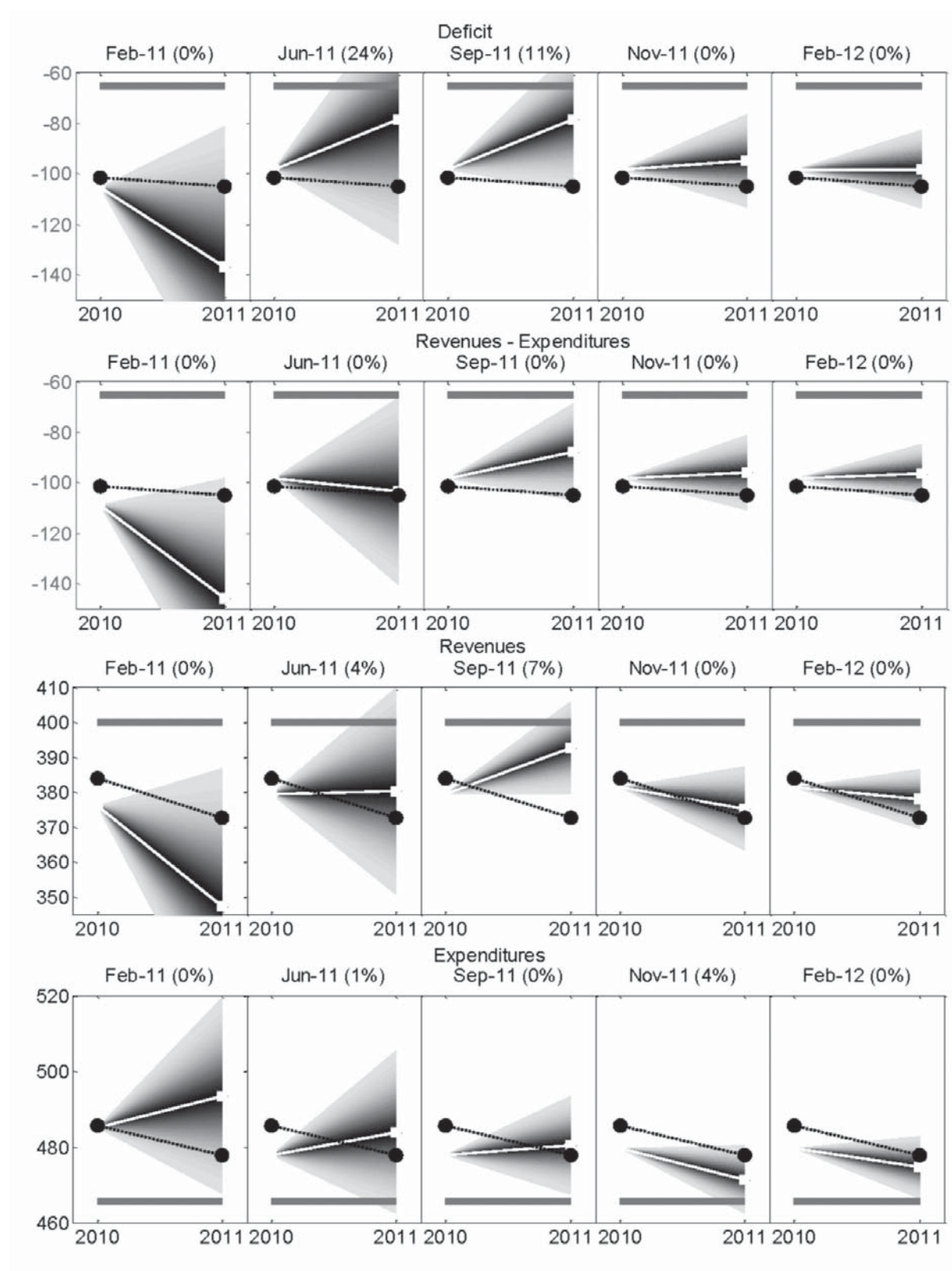


Figure 2. Forecasts and fancharts with the limit of 99% confidence^{a,b}

^a Dots: actual values; White: forecasts; Grey wide: 2011 target.

^b Figures in parenthesis are probabilities of meeting the annual targets at each date.

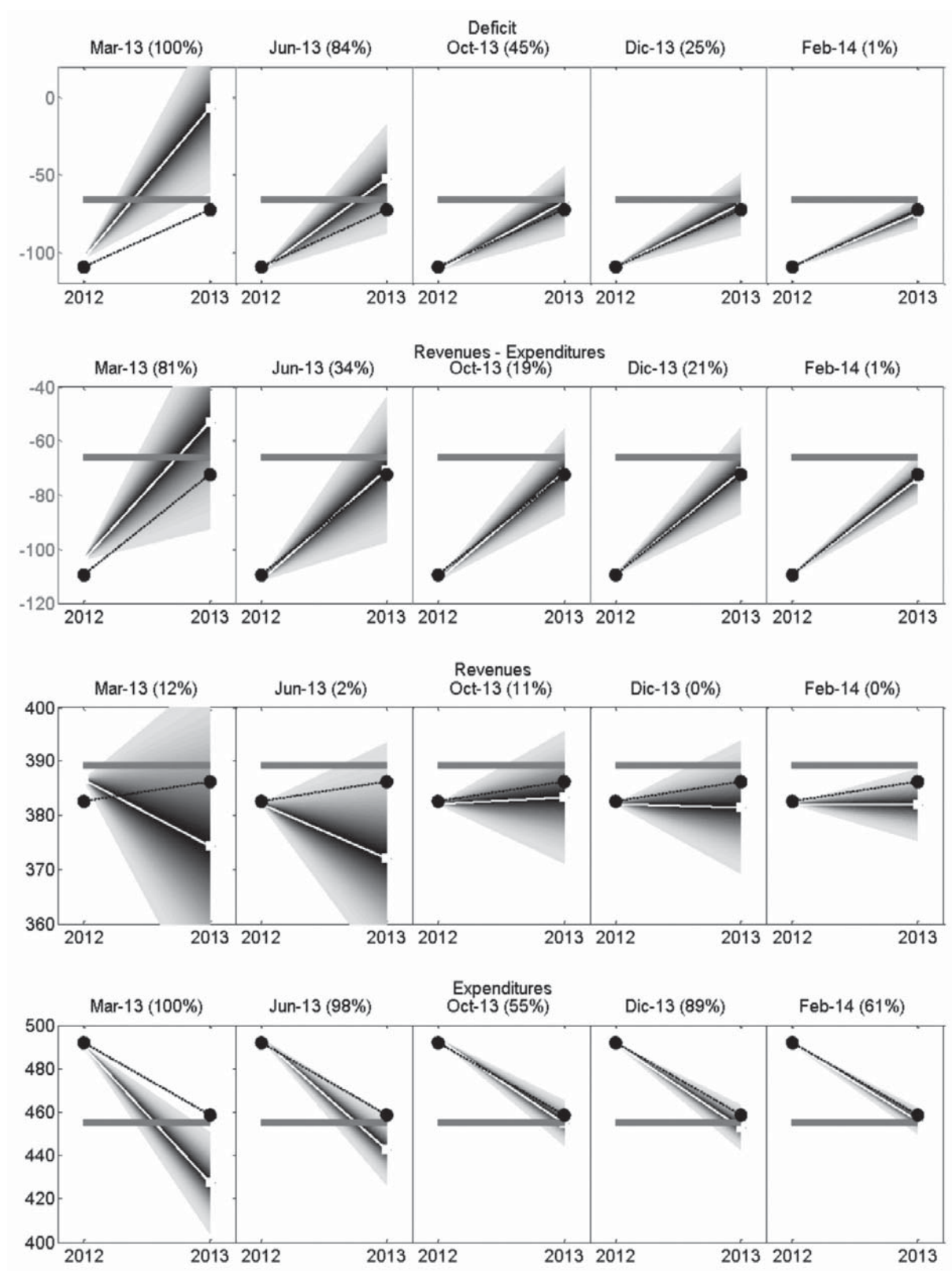


Figure 3. Forecasts and fancharts with the limit of 99% confidence^{a,b}

^a Dots: actual values; White: forecasts; Grey wide: 2011 target.

^b Figures in parenthesis are probabilities of meeting the annual targets at each date.

Table 7
THE EMPIRICAL DISTRIBUTION FUNCTION OF MODELS' FORECAST ERRORS
Panel A: Intra-annual evolution

| | General government balance | | | | Total revenues | | | | Total expenditures | | | |
|-----------------|----------------------------|-----------|------------|-----------|----------------|-----------|------------|-----------|--------------------|-----------|------------|-----------|
| | All | Q1 | Q2-Q3 | Q4 | All | Q1 | Q2-Q3 | Q4 | All | Q1 | Q2-Q3 | Q4 |
| Nº Obs: | 312 | 78 | 156 | 78 | 208 | 52 | 104 | 52 | 208 | 52 | 104 | 52 |
| Distrib (%) | | | | | | | | | | | | |
| $(-\infty, -4]$ | 8,01 | 14,10 | 8,33 | 1,28 | 9,62 | 13,46 | 10,58 | 3,85 | 1,44 | 3,85 | 0,96 | 0,00 |
| $(-4, -2]$ | 8,65 | 7,69 | 10,90 | 5,13 | 5,29 | 3,85 | 6,73 | 3,85 | 8,65 | 11,54 | 7,69 | 7,69 |
| $(-2, 0]$ | 38,46 | 32,05 | 37,18 | 47,44 | 45,67 | 42,31 | 46,15 | 48,08 | 45,19 | 32,69 | 48,08 | 51,92 |
| $(0, 2]$ | 25,64 | 28,21 | 22,44 | 29,49 | 25,48 | 25,00 | 21,15 | 34,62 | 42,79 | 48,08 | 41,35 | 40,38 |
| $(2, 4]$ | 11,86 | 8,97 | 12,82 | 12,82 | 8,17 | 5,77 | 9,62 | 7,69 | 1,92 | 3,85 | 1,92 | 0,00 |
| $(4, +\infty)$ | 7,37 | 8,97 | 8,33 | 3,85 | 5,77 | 9,62 | 5,77 | 1,92 | 0,00 | 0,00 | 0,00 | 0,00 |

Panel B: Good and Bad times

| | General government balance | | | Total revenues | | | Total expenditures | | |
|-----------------|----------------------------|-----------------|-------------------|----------------|-----------------|-------------------|--------------------|-----------------|-------------------|
| | All (00-12) | Boom (00-07) | Crisis (08-12) | All (00-12) | Boom (00-07) | Crisis (08-12) | All (00-12) | Boom (00-07) | Crisis (08-12) |
| Nº Obs: | 312 | 192 | 120 | 208 | 128 | 80 | 208 | 128 | 80 |
| Distrib (%) | | | | | | | | | |
| $(-\infty, -4]$ | 8,01 | 0,52 | 20,00 | 9,62 | 0,78 | 23,75 | 1,44 | 0,00 | 3,75 |
| $(-4, -2]$ | 8,65 | 6,25 | 12,50 | 5,29 | 2,34 | 10,00 | 8,65 | 2,34 | 18,75 |
| $(-2, 0]$ | 38,46 | 52,08 | 16,67 | 45,67 | 61,72 | 20,00 | 45,19 | 41,41 | 51,25 |
| $(0, 2]$ | 25,64 | 34,90 | 10,83 | 25,48 | 30,47 | 17,50 | 42,79 | 53,91 | 25,00 |
| $(2, 4]$ | 11,86 | 6,25 | 20,83 | 8,17 | 3,13 | 16,25 | 1,92 | 2,34 | 1,25 |
| $(4, +\infty)$ | 7,37 | 0,00 | 19,17 | 5,77 | 1,56 | 12,50 | 0,00 | 0,00 | 0,00 |

^a The numbers shows the percentage of observations within each interval. The range limits are expressed in percentage of GDP.

^b All the models discussed previously have been included. See footnotes to Table 5.

7. Conclusions and policy discussion

In this paper we present a comprehensive fiscal forecasting system, based on all short-term fiscal data available for the Spanish case. Our system is made of a suite of models, with different levels of disaggregation (bottom-up vs top-down; general government vs sub-sectors) suitable for the automatic processing of the large amount of monthly/quarterly fiscal data published nowadays by Spanish statistical authorities.

Beyond presenting the tools as such, in this paper we show some example of its potential applications for real-time monitoring of public finances. In particular, we show how the combination of models provides extremely accurate signals when information pertaining to the first half of the year is available, both in quantitative and qualitative terms. Surprisingly enough, the models contain information that seemed not to have been factored into European

Commission fiscal forecasts, that are among the best performers within the set of international organizations, and supposedly incorporate not only past data, but also forward-looking information on approved, but not yet implemented, policy measures. In addition, we provide truly real-time applications for the fiscal years 2011 and 2013, showing how the models adapted to the new information and presented reliable signals for end-of-the-year fiscal outcomes. Thus, we claim that our tools might potentially be instrumental for ex-ante detection of risks to official projections, and thus can help in reducing the ex-post reputational costs of budgetary deviations.

From a policy-making point of view, we also claim that official monitoring bodies could incorporate in their toolkit to evaluate regular adherence to targets more formal elements, of the kind presented in our paper, in order to move the standard evaluation procedure beyond the extant, more “legalistic” approach. In addition, presenting model-based results, or uncertainty tests around government targets, as the ones shown in our real-time exercise, may be helpful to convey to the public risks surrounding fiscal projections (on this subject see also Clark *et al.*, 2013). Incorporating these type of elements may help in improving communication policies as regards sources of risks of (ex-ante) compliance with budgetary targets as well as reasons for (ex-post) budgetary deviations had them occurred.

Appendix: econometric methodology

All the models developed in our paper fit with the following general discussion. The description follows quite closely Pedregal and Pérez (2010). The starting point of the modeling approach is to consider a multivariate Unobserved Components Model known as the Basic Structural Model (Harvey, 1989). A given time series is decomposed into unobserved components which are meaningful from an economic point of view (trend, T_t , seasonal, S_t , and irregular, e_t). Equation (4) displays a general form, where t is a time sub-index measured in quarters, z_t denotes the variable in ESA terms expressed at an annual and quarterly sampling interval (depending on availability) for our objective time series, and u_t represents the vector of quarterly indicators.

$$\begin{bmatrix} z_t \\ u_t \end{bmatrix} = T_t + S_t + e_t \quad (4)$$

The general consensus in this type of multivariate models in order to enable identifiability is to build SUTSE models (Seemingly Unrelated Structural Time Series). This means that components of the same type interact among them for different time series, but are independent of any of the components of different types. In addition, statistical relations are only allowed through the covariance structure of the vector noises, but never through the system matrices directly. This allows that, trends of different time series may relate to each other, but all of them are independent of both the seasonal and irregular components. The full model is a standard BSM that may be written in State-Space form as (see Harvey, 1989).

$$x_t = \Phi X_{t-1} + E w_t \quad (5)$$

$$\begin{bmatrix} z_t \\ u_t \end{bmatrix} = \begin{bmatrix} H \\ H^u \end{bmatrix} x_t + \begin{bmatrix} \epsilon_t \\ v_t \end{bmatrix} \quad (6)$$

where $\epsilon_t \sim N(0, \Sigma_E)$ and $v_t \sim N(0, \Sigma_v)$. The system matrices Φ , E , H and H^u in equations (5)-(6) include the particular definitions of the components and all the vector noises have the usual Gaussian properties with zero mean and constant covariance matrices (ϵ_t and v_t are correlated among them, but both are independent of w_t). The particular structure of the covariance matrices of the observed and transition noises defines the structures of correlations among the components across output variables. The mixture of frequencies, and the estimation of models at the quarterly frequency, implies combining variables that at the quarterly frequency can be considered as stocks with those being pure flows. This may be achieved in two steps: i) rewriting system (5)-(6) in a convenient way and ii) adding an accumulator variable to such system as defined below.

It is easy to check that model (5)-(6) is exactly equivalent to model (7)-(8), by extending the state vector with the output variables.

$$\begin{bmatrix} z_t \\ x_t \end{bmatrix} = \begin{bmatrix} 0 & H\Phi \\ 0 & \Phi \end{bmatrix} \begin{bmatrix} z_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} I & HE \\ 0 & E \end{bmatrix} \begin{bmatrix} \epsilon_t \\ w_t \end{bmatrix} \quad (7)$$

$$\begin{bmatrix} z_t \\ u_t \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & H^u \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} v_t \quad (8)$$

In order to take into account time aggregation, i.e. the fact that the output variable and the inputs are sampled at different frequencies (annual vs quarterly), an accumulator variable in the fastest sampling interval ought to be defined as in equation (9). Beware that the sampling interval of the output variable does not need to be regular.

$$C_t = \begin{cases} 0, & t = \text{one sample after every non missing data point} \\ 1, & \text{otherwise} \end{cases} \quad (9)$$

In this way system (7)-(8) becomes (10)-(11). Beware that by setting $C_t = 0$ we return actually to the previous system.

$$\begin{bmatrix} z_t \\ x_t \end{bmatrix} = \begin{bmatrix} C_t \otimes IH\Phi \\ 0 & \Phi \end{bmatrix} \begin{bmatrix} z_{t-1} \\ x_{t-1} \end{bmatrix} + \begin{bmatrix} I & HE \\ 0 & E \end{bmatrix} \begin{bmatrix} \epsilon_t \\ w_t \end{bmatrix} \quad (10)$$

$$\begin{bmatrix} z_t \\ u_t \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & Hu \end{bmatrix} \begin{bmatrix} z_t \\ x_t \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} v_t \quad (11)$$

Given the structure of the system and the information available, the Kalman Filter and Fixed Interval Smoother algorithms provide an optimal estimation of states. Maximum likelihood in the time domain provides optimal estimates of the unknown system matrices, which in the present context are just covariance matrices of all the vector noises involved in the model. The use of the models selected and the estimation procedures described in the previous paragraph, allows the estimation of models with unbalanced data sets, i.e. input variables with different sample lengths. This is a feature of relevance for the construction of the database at hand, given occasional differences in temporal coverage of indicators.

In our case, particular empirical specifications for each variable will be considered in the light of the available information (fiscal indicators). Let us provide some examples. For instance, for the case of the individual model for total government revenues, z comprises total government revenues in National Accounts terms, a variable that is available at the annual frequency from 1986-1999 and at the quarterly frequency from 2000Q1-2012Q4, while u is a matrix composed of three series (available at the quarterly frequency for the whole sample period): (i) a proxy to general government total revenues in public accounts (cash) terms; (ii) Central government total revenues and (iii) Social Security (SSS+SPEE) sector's total revenues. Models with more than one National Accounts variable also fit within the general formulation of this Section. In those cases, z includes several variables and thus u would have been a matrix with indicators by blocks for each component of z . For example, in the case of the model that jointly estimates and forecasts total revenues (TOR) and total expenditures (TOE) $z = \{\text{TOR}, \text{TOE}\}$.

Notes

1. 13 November 2013, Reuters on-line, "Spain sets up fiscal watchdog, doubts linger on independence".
2. In particular the legal provisions on budgetary surveillance procedures included in the new Budgetary Stability Law (May 2012), and other comptroller institutions like the IGAE (State Accounts Comptroller) or the Court of Exchequer ("Tribunal de Cuentas"). At the end of 2013 the Spanish Government also created an "Independent Fiscal Authority" that should help in the monitoring process of budgetary plans.
3. Other approaches for modeling data at different sampling intervals are the methods based on regression techniques (Chow and Lin, 1971, Guerrero, 2003), the MIDAS (MIXed DATA Sampling) approach (see Ghysels, Santa-Clara & Valkanov, 2004, Clements and Galvão, 2007), the state space approaches of Liu and Hall (2001) and Mariano and Murusawa (2003), or the ARMA model model with missing observations of Hyung and Granger (2008).
4. Along the same lines see also the discussion of Leal, Pedregal and Pérez (2010).
5. We focus in our paper on nominal budgetary variables, despite the fact that typically fiscal targets are set as percentages of GDP. We decided to take this avenue, along the lines of Pérez and Pedregal (2010), because forecasting GDP, being the key macroeconomic variable, is a science in itself. That is why we decided to focus on nominal fiscal values and not ratios-to-GDP, as any error/signal to be taken from our toolkit would mix genuine macro and fiscal errors that could not be easily disentangled. This said, we see our "fiscal forecasting models" for nominal quantities as an input to an overall macroeconomic forecasting exercise, in which GDP is forecasted, as it is the standard practice in the field of fiscal forecasting (see e.g. Leal *et al.*, 2008).
6. For applications with euro area aggregate data and/or pools of euro area countries' data see Pérez (2007), Silvestrini *et al.* (2008), Onorante *et al.* (2009), Pedregal and Pérez (2010), Paredes *et al.* (2014), Asimakopoulou *et al.* (2013)

7. The standard practice for factoring-in new intra-annual fiscal information on revised annual fiscal targets and projections is via informed, judgemental add-in factors (Leal *et al.*, 2008). The fact that budgetary projections are prepared in annual terms, given an annual budgetary cycle, and the discretionary nature of many government measures set up for the entire year, have traditionally limited the role of high-frequency fiscal data for monitoring annual budgetary targets in the course of the year.
8. See also Leal and Pérez (2005).
9. As regards the latter, a database of monthly and quarterly fiscal indicators is available upon request (see also de Castro *et al.*, 2014 for details on the input database) as well as a fully-documented MATLAB toolbox that uploads the data, runs all the models and provides standardized output files. All in all, some 20 different models are run over a dataset including more than 200 fiscal variables, and this is done automatically in a few hours of computing time.

References

- Annett, A. (2006), “Enforcement and the Stability and Growth Pact: How Fiscal Policy Did and Did Not Change Under Europe’s Fiscal Framework”, IMF Working Paper 2006/116.
- Argimón, I. and Martí, F. (2006), “Las finanzas públicas de las Comunidades Autónomas: presupuestos y liquidaciones”, *Boletín Económico del Banco de España*, December.
- Artis, M. and Marcellino, M. (2001), “Fiscal forecasting: the track record of the IMF, the OECD and the EC”, *The Econometrics Journal*, 4: S20-S36.
- Asimakopoulou, S.; Paredes, J. and Warmedinger, T. (2013), “Forecasting fiscal time series using mixed frequency data”, Working Paper Series of the European Central Bank 1550.
- Chow, G. C. and Lin, A. (1971), “Best Linear Interpolation, distribution and extrapolation of time series by related series”, *Review of Economics and Statistics*, 53: 372-375.
- Clark, J.; Gibbons, C.; Morrissey, S.; Pooley, J.; Pye, E.; Wilcox, R. and Willard, L. (2013), “Estimates of uncertainty around budget forecasts”, *Australian Treasury Working Paper* 2013-04.
- Clements, M. P. and Galvão, A. B. (2007), “Macroeconomic forecasting with mixed frequency data: Forecasting US output growth”. Working Paper 616, Department of Economics, University of London.
- de Castro, F.; Martí, F.; Montesinos, A.; Pérez, J. J. and Sánchez, A. J. (2014), “Fiscal policies in Spain: main stylized facts revisited”, Banco de España. Working Paper, N° 1408.
- Edo, V. (2012), “Los programas presupuestarios del Estado en España: metodología y análisis de las desviaciones presupuestarias 1990-2009”, *Revista Auditoría Pública*, 56.
- European Commission (2013), “Building a strengthened fiscal framework in the European Union: a guide to the Stability and Growth Pact”, *European Economy Occasional Papers*, 150.
- Fernández-Caballero, L.; Pedregal, D. J. and Pérez, J. J. (2012), “Monitoring sub-central government spending in Spain”, *Hacienda Pública Española/Review of Public Economics*, 202: 77-104.
- Frankel, J. A. and Schreger, J. (2013), “Over-optimistic official forecasts and fiscal rules in the Eurozone”, *Review of World Economics*, forthcoming.
- Ghysels, E.; Santa-Clara, P. and Valkanov, R. (2004), “Predicting volatility: Getting the most out of return data sampled at different frequencies”, *Journal of Econometrics*.

- Gómez, N. and Guerrero, V. M. (2006), 'Restricted forecasting with VAR models: An analysis of a test for joint compatibility between restrictions and forecasts', *International Journal of Forecasting*, 22: 751-770.
- Gómez, V. and Maravall, A. (1996), "Programs TRAMO (Time series Regression with Arima noise, Missing observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series). Instructions for the User", Banco de España Working Paper No. 9628.
- Guerrero, V. M. (2003), "Monthly disaggregation of a quarterly time series and forecasts of its unobservable monthly values", *Journal of Official Statistics*, 19: 215-235.
- Hameed, F. (2005), "Fiscal transparency and economic outcomes", IMF Working Paper n. 05/225. Washington: International Monetary Fund.
- Hannon, P. (2014), "Difficulties in Forecasting Borrowing Requirements", The Wall Street Journal, 10 January 2014, available at: "blogs.wsj.com/moneybeat/2014/01/10/difficulties-in-forecasting-borrowing-requirements/"
- Harvey, A. (1989), *Forecasting Structural Time Series Models and the Kalman Filter*. Cambridge University Press. UK.
- Harvey, A. and Chung, C. (2000), "Estimating the underlying change in unemployment in the UK", *Journal of the Royal Statistical Society, Series A*: 163: 303-339.
- Hyung, N. and Granger, C. J. (2008), "Linking series generated at different frequencies", *Journal of Forecasting*, 27: 95-108.
- Hernández de Cos, P. and Pérez, J. J. (2013a), "Sub-national public debt in Spain: political economy issues and the role of fiscal rules and decentralization", pp. 188-216 in "Fiscal relations across government levels in times of crisis making compatible fiscal decentralization and budgetary discipline", European Commission Economics Papers, 501. Brussels.
- Hernández de Cos, P. and Pérez, J. J. (2013b), "The new budgetary stability law", Economic Bulletin, Bank of Spain, April.
- International Monetary Fund (2010), "Fiscal transparency, accountability and risk", International Monetary Fund Policy Paper 080712. Washington D.C.
- Jonung, L. and Larch, M. (2006), "Fiscal policy in the EU: are official output forecasts biased?", *Economic Policy*: 491-534, July.
- Kanda, D. (2002), "Assessing Monthly Progress Toward Annual Fiscal Revenue Targets", IMF WP/02/116.
- Keereman, F. (1999), "The track record of Commission Forecasts", European Commission Economic Papers, 137.
- Leal, A. and López Laborda, J. (2013), "Un estudio de los factores determinantes de las desviaciones presupuestarias de las Comunidades Autónomas en el periodo 2003-2012", *Documentos de Trabajo*, 735/2013, Fundación de las Cajas de Ahorros.
- Leal, T.; Pérez, J. J.; Tujula, M. and Vidal, J.-P. (2008), "Fiscal forecasting: lessons from the literature and challenges", *Fiscal Studies*, 29: 347-386.
- Leal, T. and Pérez, J. J. (2009), "Un sistema ARIMA con agregación temporal para la previsión y el seguimiento del déficit del Estado", *Hacienda Pública Española/Review of Public Economics*, 190: 27-58.

- Leal, T. and Pérez, J. J. (2005), “Seguimiento de objetivos fiscales anuales: una aplicación con datos regionales”, *Investigaciones Regionales*, 7: pp. 53-67, November.
- Leal, T. and Pérez, J. J. (2011), “Análisis de las desviaciones presupuestarias aplicado al caso del presupuesto del Estado”, *Estudios de Economía Aplicada*, 29-3: 1-14.
- Leal, T.; Pedregal, D. J. and Pérez, J. J. (2010), “Short-term monitoring of the Spanish government balance”, *SERIEs*.
- Liu, H. and Hall, S. G. (2001), “Creating high-frequency national accounts with state space modelling: A Monte Carlo experiment”, *Journal of Forecasting*, 20: 441-449.
- Mariano, R. S. and Murasawa, Y. (2003), “A new coincident index of business cycles based on monthly and quarterly series”. *Journal of Applied Econometrics*, 18: 427-443.
- Merola, R. and Pérez, J. J. (2013), “Fiscal forecast errors: governments versus independent agencies?”, *European Journal of Political Economy*, 23: 285-299.
- Moauero, F. and Savio, G. (2005), “Temporal Disaggregation Using Multivariate Structural Time Series Models”, *Econometrics Journal*, 8: 214-234.
- Moulin, L. and Wierds, P. (2006), How Credible are Multiannual Budgetary Plans in the EU?, *Fiscal Indicators*, Banca d'Italia.
- Onorante, L.; Pedregal, D. J.; Pérez, J. J. and Signorini, S. (2009), “The usefulness of infra-annual government cash budgetary data for fiscal forecasting in the euro area”, *Journal of Policy Modeling*, 32: 98-119.
- Paredes, J.; Pedregal, D. J. and Pérez, J. J. (2014), “Fiscal policy analysis in the euro area: expanding the toolkit”, *Journal of Policy Modeling*, 36: 800-823.
- Pedregal, D. J. and Pérez, J. J. (2010), “Should quarterly government finance statistics be used for fiscal surveillance in Europe?”, *International Journal of Forecasting*, 26: 794-807.
- Pérez, J. J. (2007), “Leading Indicators for Euro Area Government Deficits”, *International Journal of Forecasting*, 23: 259-275.
- Pina, A. and Venes, N. (2011), “The political economy of EDP fiscal forecasts: an empirical assessment”, *European Journal of Political Economy*, 27: 534-546.
- Pons, J. and Solé, A. (2001), “Racionalidad de las previsiones presupuestarias y comportamiento estratégico. El caso español”, *Hacienda Pública Española*, 157 (2): 197-232.
- Proietti, T. and Moauero, F. (2006), “Dynamic factor analysis with non-linear temporal aggregation constraints”. *Journal of the Royal Statistical Society Series C, Royal Statistical Society*, 55: 281-300.
- Silvestrini, A.; Salto, M.; Moulin, L. and Veredas, D. (2008), “Monitoring and forecasting annual public finance deficit every month: the case of France”. *Empirical Economics*, 34: 493-524.
- %item Solé-Ollé and Bosch (2007)
- Strauch, R.; Hallerberg, M. and von Hagen, J. (2004), “Budgetary Forecasts in Europe - The Track Record of Stability and Convergence Programmes”, European Central Bank WP 307.
- von Hagen, J. (2010), “Sticking to fiscal plans: the role of institutions”, *Public Choice*, 144: 487-503.

Resumen

En este artículo desarrollamos un sistema exhaustivo para la predicción fiscal a corto plazo, útil para supervisar en tiempo real las necesidades de financiación de las AAPP en España, un país que ha estado situado en los últimos años en el centro de la reciente crisis de deuda Europea, por las desviaciones observadas respecto a los objetivos fijados para el déficit público. El sistema está compuesto por una variedad de modelos, con distintos niveles de desagregación y resulta adecuado para el proceso automático del gran volumen de información presupuestaria mensual y trimestral publicadas a día de hoy por las autoridades estadísticas españolas. Nuestra herramienta es útil para la detección ex ante de riesgos de cumplimiento y, consecuentemente, puede ayudar a reducir el coste reputacional derivado de las desviaciones presupuestarias registradas a posteriori. Sobre la base de nuestros resultados, discutimos cómo los analistas pueden expandir, por un lado, las herramientas disponibles para evaluar el cumplimiento de los objetivos (yendo más allá de las obligaciones meramente formales) y, por el otro, la política de comunicación de los riesgos ex ante de cumplimiento de los objetivos presupuestarios.

Palabras clave: Responsabilidad del gobierno, transparencia, predicción fiscal.

Clasificación JEL: E62, E65, H6, C3, C82.