

Der Open-Access-Publikationsserver der ZBW – Leibniz-Informationzentrum Wirtschaft  
*The Open Access Publication Server of the ZBW – Leibniz Information Centre for Economics*

Leal, Teresa; Pedregal, Diego J.; Pérez, Javier J.

**Article**

## Short-term monitoring of the Spanish government balance

SERIEs - Journal of the Spanish Economic Association

**Provided in Cooperation with:**  
Spanish Economic Association

Suggested Citation: Leal, Teresa; Pedregal, Diego J.; Pérez, Javier J. (2011) : Short-term monitoring of the Spanish government balance, SERIEs - Journal of the Spanish Economic Association, ISSN 1869-4195, Springer, Heidelberg, Vol. 2, Iss. 1, pp. 97-119, <http://dx.doi.org/10.1007/s13209-010-0018-3>

This Version is available at:  
<http://hdl.handle.net/10419/77741>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

## Short-term monitoring of the Spanish government balance

Teresa Leal · Diego J. Pedregal · Javier J. Pérez

Received: 23 July 2009 / Accepted: 26 October 2009 / Published online: 21 January 2010  
© Spanish Economic Association and Fundación SEPI 2010

**Abstract** We construct multivariate, state-space mixed-frequencies models for the main components of the Spanish General Government sector made up of blocks for each one of its subsectors: Central Government, Social Security and aggregate of Regional and Local government sectors. Each block is modelled through its total revenue and expenditure categories, and encompasses a number of indicators, depending on data availability. The mixed-frequencies approach is particularly relevant for the case of Spain, given its institutional set-up and the specific data availability for the different subsectors. All in all, we provide models detailed enough in coverage, while at the same time manageable, to be used: (1) for real-time monitoring of fiscal policies with a focus on quarterly developments of the General Government sector, (2) for the monitoring of general government sub-sectors for which intra-annual data coverage is limited (Regional and Local governments), (3) to bridge (translate) into National Accounts available monthly information for the subsectors of the general government.

**Keywords** Fiscal forecasting · Fiscal policies · Mixed frequency data · Kalman filter

**JEL Classification** C53 · E6 · H6

---

T. Leal  
University of Huelva, Huelva, Spain  
e-mail: mtleal@uhu.es

D. J. Pedregal  
U. Castilla-La Mancha, Ciudad Real, Spain  
e-mail: diego.pedregal@uclm.es

J. J. Pérez (✉)  
Banco de España, Madrid, Spain  
e-mail: javierperez@bde.es

## 1 Introduction

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 practitioners. The fact that budgetary projections are prepared in annual terms, given an annual budgetary cycle, in the framework of annual models, 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. The relative scarcity of intra-annual data for the desired, policy-relevant government variables (typically referred to the general government sector) has contributed to this situation as well. Thus, 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).

In this framework, official annual fiscal targets and projections tend to display well-documented political biases and/or large forecast errors (see Strauch et al. 2004; Moulin and Wierts 2006; Annett 2006; Pina and Venes 2008; Jonung and Larch 2006; Leal et al. 2008). In addition to political biases, fiscal forecast errors might be due to a number of factors, in particular policy errors, owing to the implementation of new fiscal policy measures, not yet announced by the forecast cut-off date, and economic errors due to inaccurate forecasts of the macroeconomic variables that underlie budgetary projections.

Even without judging the determinants of fiscal forecasts, some recent episodes show large forecast errors linked to official annual fiscal projections. Take for example the case of the Spanish General Government deficit in 2008. In this year, according to the Spring 2009 EDP notification, the general government recorded a deficit of 3.8% of GDP. By June 2008 most national and international institutions still projected a budgetary surplus for 2008, and only some institutions timidly turned their estimates to small deficits for the whole year after the summer. Nevertheless, as late as October 2008 the government still estimated a deficit of 1.5% of GDP, slightly above the 1.6% deficit projected by the European Commission around the same date (see EC 2008). The same estimate for 2008 was kept as a reference by the government in the budget law for 2009 that passed parliamentary approval at the end of December. At the beginning of January, though, in the framework of the updated Stability Programme for 2008–2012, the government provided an estimated deficit for 2008 of 3.4% of GDP, close to the final figure. A natural question arises: was the sharp, unanticipated revision of the estimated deficit by the government indeed unpredictable or was it motivated by some strategic behaviour?—as the literature on politically-motivated fiscal forecasts might have suggested. In a related fashion, and more relevant for the aim of this paper: given the available information at the time, would it have been possible for an independent analyst to detect the fiscal deterioration in advance, at least in the course of the last part of 2008?

On the development of early-warning tools for fiscal variables, a recent strand of the literature has shown that intra-annual fiscal data, when modelled appropriately, contains extremely valuable and useful information for forecasting annual fiscal aggregates, enabling earlier detection of episodes of fiscal deterioration (or improvement)

than traditional methods (Pérez 2007; Silvestrini et al. 2008; Onorante et al. 2009; Pedregal and Pérez 2009; Leal and Pérez 2009).

The present paper aims at contributing to this branch of the literature in two directions. Firstly, it focuses with high detail on one single country (Spain) what enables the development of a model for the General Government sector (GG) made up of blocks for each one of its subsectors: the Central Government sector (CG), the Social Security sector (SS) and the aggregate of the Regional and Local government sectors (LG). To exploit the existing accounting identity between GG and its subsectors ( $GG = CG + SS + LG$ ), the variables total revenue, total expenditure and deficit are modelled in our baseline model in such a way that  $Z_{GG} = Z_{CG} + Z_{SS} + Z_{LG}$ , where  $Z$  stands for any of the fiscal variables mentioned above. In addition, each subsector is modelled through its total revenue and expenditure categories, and encompasses a number of idiosyncratic indicators, depending on data availability. The extant literature either focused on the State subsector of the general government, or directly on the aggregate of the general government for some representative, headlines fiscal variables.

Secondly, along the lines of the models by Pedregal and Pérez (2009) for euro area fiscal aggregates, the models presented in the current paper make use of annual, quarterly and monthly ESA95 fiscal figures and other relevant monthly fiscal indicators. An optimal way to use these data is to build a single model that relates data at all frequencies. Thus, we construct multivariate, state-space mixed-frequencies models along the lines of Harvey and Chung (2000), Moauro and Savio (2005), Proietti and Moauro (2006), and Frale and Veredas (2008).<sup>1</sup> This is particularly relevant for the case of Spain, given important differences in data availability for the different subsectors of the general government. Quarterly ESA95 (European System of National Accounts) figures are available only for the general government sector as a whole, for the period starting in 1995, but not for any of its individual subsectors. For the subsectors as defined in National Accounts only annual figures are available for the sectors. A partial exception is the central Government sector, given that monthly ESA95 figures are available for its main component (the State sector) that comprises almost 90% of the total. As regards the Social Security sector, there is monthly information, though on a cash basis, for several variables. As regards the Regional and Local government sectors, the availability of intra-annual information is extremely limited.

All in all, we provide models detailed enough in coverage, while at the same time manageable, to be used: (1) for real-time monitoring of fiscal policies with a focus on quarterly developments of general government figures, (2) as an input for the preparation of annual fiscal projections, (3) to bridge (translate) into National Accounts available monthly information for the different subsectors of the general government, (4) to make inferences on short-term developments of Regional and Local governments, given the fact that quarterly information for the aggregate of the General Government

---

<sup>1</sup> 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 et al. 2006; Clements and Galvão 2008), the state space approaches of Liu and Hall (2001) and Mariano and Murasawa (2003), or the ARMA model with missing observations of Hyung and Granger (2008).

sector is used in conjunction with monthly information for the Central and the Social Security sectors.

The models could be easily applied to other euro area countries; in particular those with a similar institutional setup and at least the same data coverage by subsectors (like Germany or Italy). For the case of Spain, this paper has to be seen as a first attempt to develop a short-term fiscal model for the General Government; subsequent refinements of the model, especially by further enlarging the data coverage by subsectors, are warranted. On different grounds, 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 models presented in this paper.

The rest of the paper is organised as follows. Section 2 describes the data used. Section 3 describes the models, Sect. 4 presents the main results of the paper, and Sect. 5 concludes.

## 2 Database description

Table 1 shows the dataset employed in the paper. It comprises a total amount of 32 time series, taken from different official providers of statistics (IGAE, INE, BDSICE), and covers the period 1985–2008. The data covers the General Government sector and its subsectors. Part of the dataset is in line with ESA95 standards, while another part follows public accounts (cash) accounting rules.

As regards direct data for the General Government sector (GG), we use annual non-financial total revenue and expenditure from 1985 to 2008, as well as quarterly non-financial total revenue and expenditure, from 1995Q1 to 2008Q4. It is worth noticing that, surprisingly enough, the quarterly information is not available for the sub-sectors comprising GG. All series are in ESA95 terms.<sup>2</sup>

By far, the Central Government sector (CG) is the subsector with the best data coverage in terms of adherence to ESA standards. We use annual non-financial total revenue and expenditure from 1985 to 2008, as well as monthly figures for most categories of the State Sector (that comprises more than 90% of the total CG sector) from 1985 to 2008. In particular, we use on the revenue side: direct taxes, VAT, other indirect taxes, and other revenues (so that the sum of the four items adds up to the total revenues of the State sector). On the expenditure side: “government consumption” (compensation of employees plus intermediate consumption), capital expenditure, interest payments, and other expenditures (so that the sum of the four items adds up to the total of expenditures of the State sector).

For the Social Security sector (SS) we also use annual non-financial total revenue and expenditure from 1985 to 2008 in ESA95 terms. The Spanish SS sector is

---

<sup>2</sup> Annual ESA95 series for the GG sector and all its subsectors are available from the IGAE (Badespe database). Within the period 1985–2008 the changeover from ESA79 to ESA95 was accounted for by applying growth rates of the variables in ESA79 terms (first part of the sample) to the levels of the variables in ESA95 (last part of the sample).

**Table 1** Data sources

Non-financial series	ESA95 coverage <sup>a</sup>	Periodicity	Sample period	Basis	Source <sup>b</sup>	Units
General Government total revenue	S.13 GG	Annual	1985–2008	ESA95	INE (BADESPE database)	Mrd EUR
General Government total expenditure						
Central Government total revenue	S.1311 CG					
Central Government total expenditure						
Social Security total revenue	S.1314 SS					
Social Security total expenditure						
General Government total revenue	S.13 GG	Quarterly	1995Q1–2008Q4	ESA95	ECB and Eurostat	
General Government total expenditure						
Direct taxes	S.1311 subsector (State)	Monthly	1985:1–2008:12	ESA95	IGAE and INE	
VAT taxes						
Other indirect taxes						
Other revenues						
Government consumption						
Capital expenditure						
Interest payments						
Other expenditures						
Total cash non-financial receipts	S.1314 subsector (Social Security System)			Cash	BDSICE database (cod.731102G)	

Table 1 continued

Non-financial series	ESA95 coverage <sup>a</sup>	Periodicity	Sample period	Basis	Source <sup>b</sup>	Units
Total contributory social benefits					BDSICE database (cod. 731210)	
Total expenditure	S.1314 subsector (Public State Employment Service)				Public State Emp. Service	
Taxes collected by the Central Government sector on behalf of the Regional government	S.1312 LG				BDSICE database (cod. 753100)	

<sup>a</sup> ESA European System Account, GG General Government, CG Central Government, LG Regional and Local governments' aggregate, SS Social Security sector

<sup>b</sup> INE Spanish National Institute, ECB European Central Bank, IGAE General Comptroller of the State Administration

basically composed of two parts. On the one hand, the “Social Security System”—covering mainly social contributions, and contributory and non-contributory social benefits—for which, monthly cash non-financial receipts and expenses are available for the period under study. On the other hand, the SS sector comprises “Other Administrations of the Social Security Funds”, which covers the “Public State Employment Service” (SPEE) and the “Social Guarantee Fund” (FOGASA). In this respect, we include in our analysis the monthly cash non-financial revenues and expenditures of the SPEE.

The Regional and Local governments (LG) are by far the sub-sectors with less available intra-annual data. In fact, no consistent time series covering a reasonable time span are available for the spending side of the LG sector. For the revenue side, available information comprises monthly cash data for the whole period regarding taxes collected by the CG sector on behalf of the LG sector (*Tributos cedidos y concertados*). Nevertheless, this source does not cover all tax revenues managed by the LG sector. As usual, we also include annual ESA95 total revenue and expenditure of the LG sector, from 1985 to 2008.

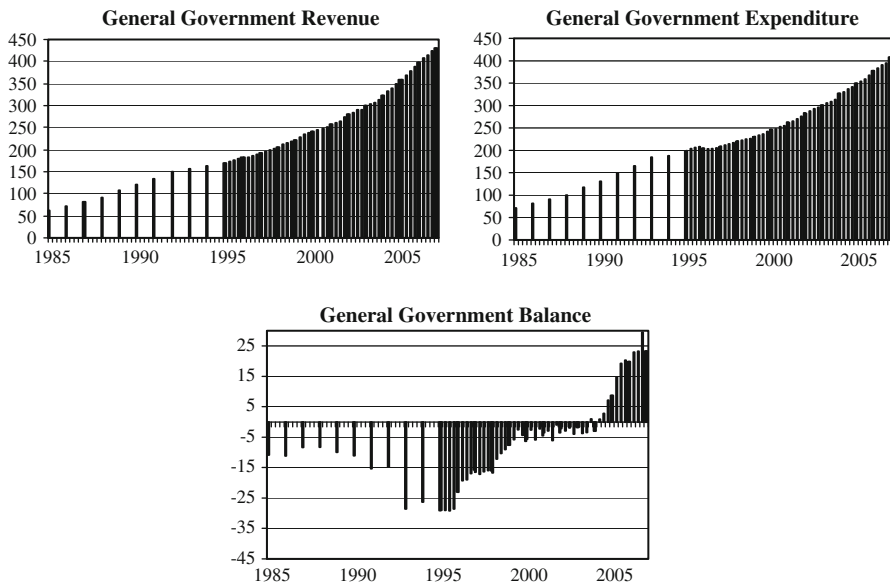
From an accounting point of view total revenues and expenditures of the GG in ESA95 terms are equal to the sum of ESA95 CG, SS and LG sectors’ total revenues and expenditures, respectively. This is a constraint that links all the annual ESA95 data described above. The database is affected by a process of decentralization by which some duties in the hands of the Central Government and Social Security sectors have been transferred to Regional and Local governments, following the particular development of the territorial structure of Spain established in the 1978 Constitution. The distribution of revenues and expenditures between the sectors has been altered by subsequent changes in the financial arrangements between them, leading to successive waves of decentralization (see [Gordo and Hernández de Cos 2001](#)). For the purpose of this paper, with a special focus on forecasting, the most influential effect arises from the 2002 financial arrangements between the Central Government and the Regional Government sectors, entering into effect in January 2002.

To address potential distortions due to this decentralisation process, we disregard a purely statistical approach consisting on balancing transfers among sectors on the basis of raw data, due to lack of sufficient information. Instead, we decide to take a simpler, econometric approach: all the affected time series are corrected beforehand with the aid of models set up in a State Space framework along the lines described in a following section. An intervention model is estimated with constraints in the intervention parameters, in a way such that the addition of the declines in CG and SS revenues and expenditures due to transfers equal the raising up of LG revenues and expenditures. Therefore, the accounting constraints mentioned earlier are respected (i.e. total GG series are untouched): revenues and expenditures in all sub-sectors are corrected so that at the end the total revenues and expenditures withdrawn from the CG and the SS sectors are exactly equal in amount to the increase in LG sector’s revenues and expenditures. In addition, the corrections are implemented in such a way that all the implied deficit (revenue minus expenditure) time series are unaffected.<sup>3</sup>

---

<sup>3</sup> Additional interventions were kept to a minimum, but some were needed on technical grounds (detected using the standard program TRAMO): additive outliers on October–November 1986 on monthly capital





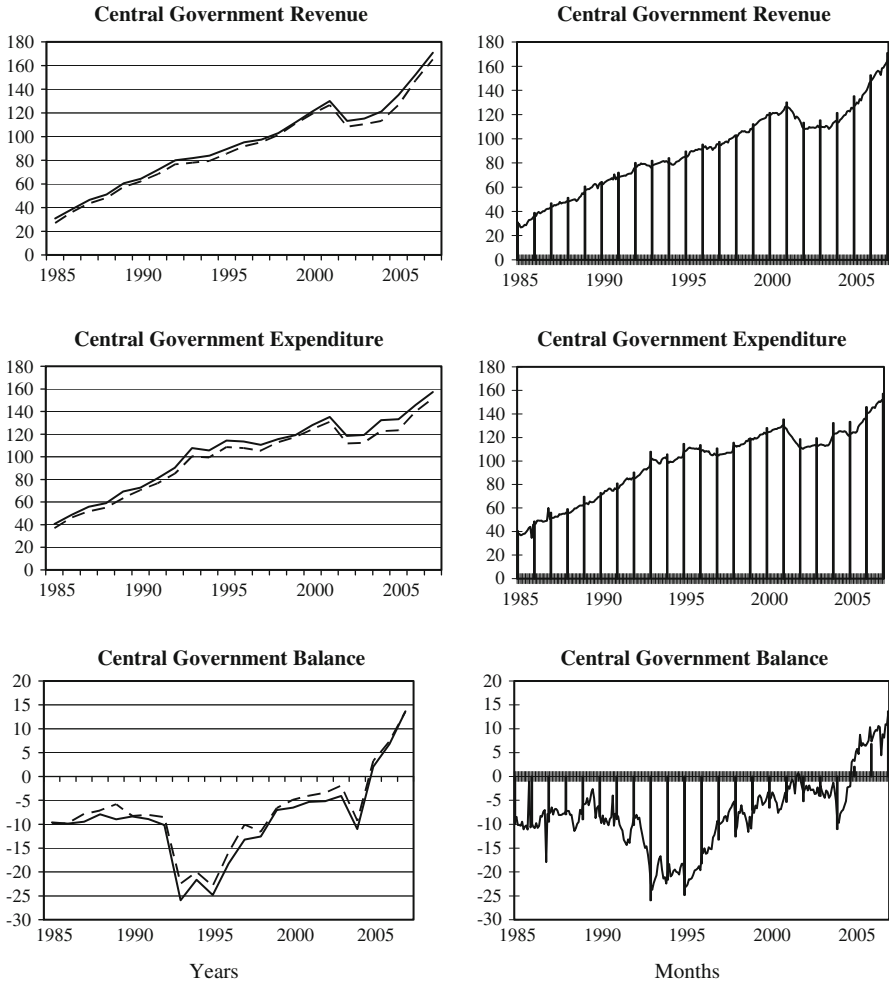
**Fig. 1** Spanish General Government revenue, expenditure and balance, 1985–2007. ESA95 annual and quarterly figures (*bars* quarterly shown as four-quarter moving sum). Billion €

Figures 1, 2, 3 and 4 show most of the variables used in the analysis: total revenues and expenditures for the GG sector and for all its subsectors (in ESA95 terms and also the available “indicators”), and also the government deficits, computed as the difference between revenues and expenditures.

Figure 1 shows the information available for the General Government: annual ESA95 series up to the year 1994 and quarterly afterwards (shown as four-quarter moving sums). Figures 2, 3 and 4 display the same information but by subsectors, and also the monthly information available in each case. Left plots in each row of these figures present the ESA95 annual series and the indicator variables (annualised monthly variables); right plots in each row show the ESA95 annual figures (bars) and monthly cash series in a monthly time axis (solid line, shown as 12-months moving sums). Figure 2 displays the expected close level correspondence between the State sector and the total Central Government sector series. Also as expected, Fig. 3 shows a quite small level gap on the revenue side between the indicators and the ESA95 aggregates for the SS sector, while total expenditures’ levels are commensurate. In Fig. 4 it is obvious that the available monthly information for the LG sector does not properly trace the evolution of the overall sector in ESA95 terms.

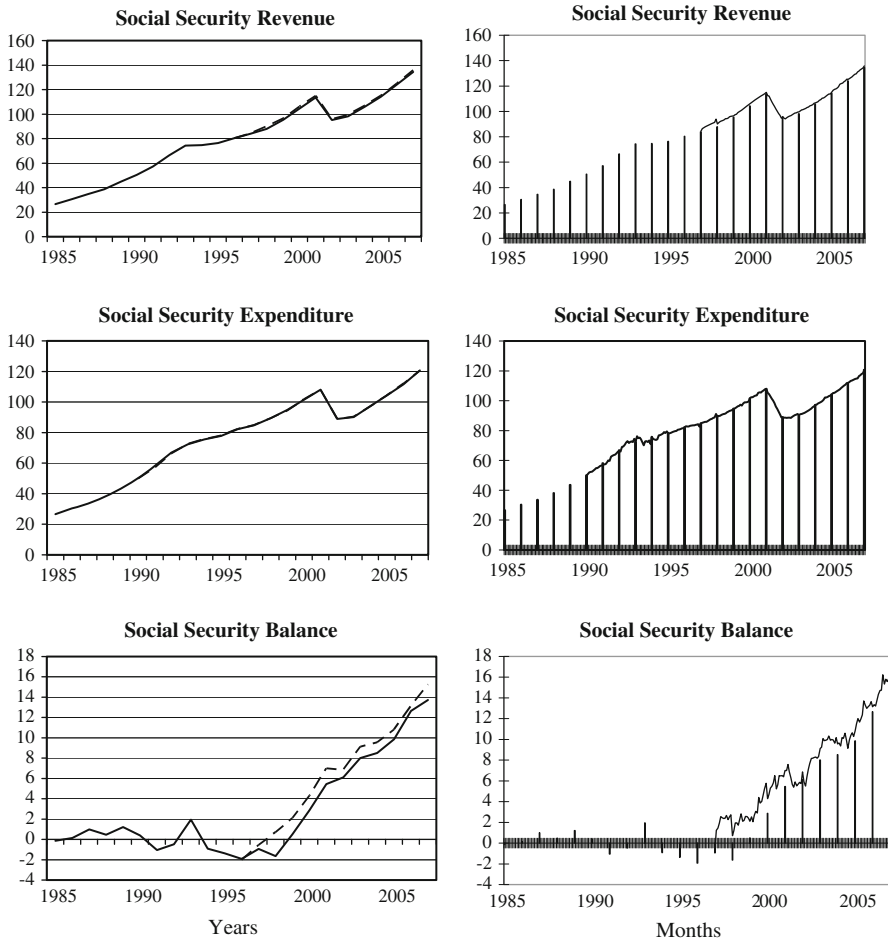
Footnote 3 continued

expenditure; additive outliers on August–September 1993 on monthly cash Social Security System’s expenditures; additive outlier on December 1998; correction of the drastic change in the variance of the series of interest payments at the beginning of the 1990s (due to changes in Treasury’s practices). Finally, the impact of one-off proceeds relative to the allocation of mobile licenses (UMTS) was removed from the ESA95 annual series and, accordingly, some adjustments were also implemented in the monthly indicators to guarantee consistency.



**Fig. 2** Spanish Central Government revenue, expenditure and balance, 1985–2007. *Left plots in each row* ESA95 annual figures (solid line) and annualised sum of 12 months for the monthly variables (dashed line). *Right plots in each row* ESA95 annual figures (bars) and sum of the four monthly indicators (solid line, shown as 12-month moving sum). Billion €

Most of the intra-annual variables included in our analysis present strong seasonal patterns. As regards government revenue variables, they largely reflect the calendar of payments of the corresponding tax categories. Income taxes are partly paid every month via withholdings on labour and capital income, but also follow an annual calendar of payments. Corporate taxes tend to follow a quarterly pattern as the bulk of the tax returns are reported quarterly by companies. Indirect tax collection, especially as regards VAT, also follows regular seasonal patterns linked to quarterly tax statements by firms and the annual refund of taxes. Government expenditures also reflect regular seasonal patterns, mainly linked to compensation of employees (payments to

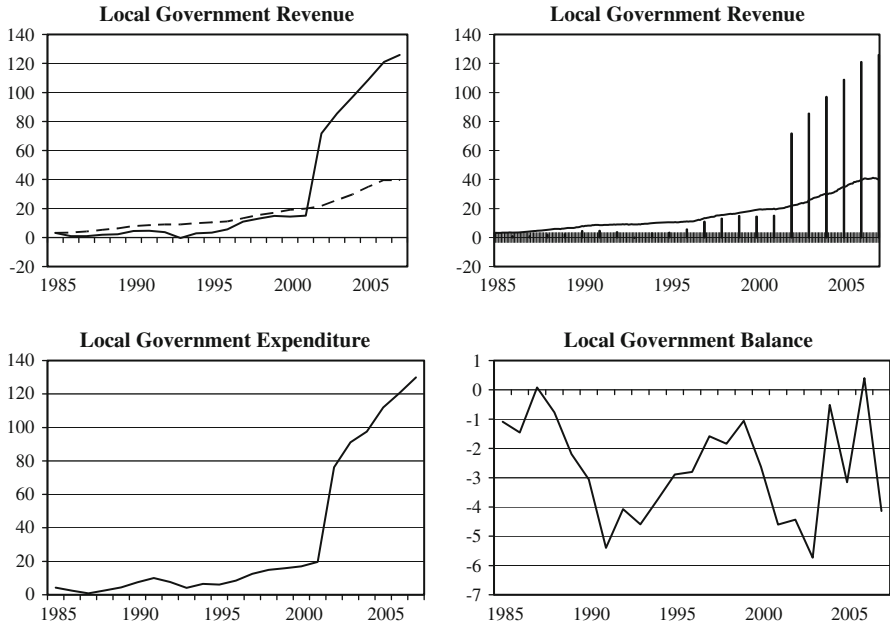


**Fig. 3** Spanish Social Security revenue, expenditure and balance, 1985–2007. *Left plots in each row* ESA95 annual figures (*solid line*) and annualised monthly cash series (*dashed line*). *Right plots in each row* ESA95 annual figures (*bars*) and monthly cash series (*solid line*, shown as 12-month moving sum). Billion €

government employees) and transfers to households (in particular old age pensions and other pensions). For these reasons the selected models, as reported in the Sect. 3 of the paper, take due account of seasonality; we model explicitly seasonal components for all the variables in our models, and estimate all the relevant seasonal parameters jointly with all other parameters of our models.

### 3 The State Space models

The mixture of frequencies, and the estimation of models at the monthly frequency, implies combining variables that at the monthly frequency can be considered as stocks with those being pure flows. An annual ESA95 series cast into the monthly frequency



**Fig. 4** Spanish Local Government revenue, expenditure and balance, 1985–2007. *Top plots* ESA95 annual figures (solid line) and annualised monthly cash series (left; dashed line); ESA95 annual figures (bars) and monthly cash series (right; solid line, shown as 12-month moving sum). *Bottom plots* annual ESA95 expenditure and deficit. Billion €

is a set of missing observations for the first months of the year (January to November) and the observed value assigned to the last month of each year (December). Theoretically, the annual ESA95 series would be obtained from a monthly ESA95 series by summation of the 12 months of a year (January to December) had them been available. In the same fashion, a quarterly ESA95 series cast at the monthly frequency encompasses missing observations for the first and the second month of each quarter, while the quarterly observation would be assigned to the last month of each quarter. In the same fashion, the quarterly ESA95 series would be obtained from a monthly ESA95 series by summation of the 3 months of each quarter had them been available.

### 3.1 Time aggregation in State Space models

Prior to the exposition of the particular formulation used in this paper (see next section), this section explains how time aggregation can be implemented in a general State Space framework (see e.g. Harvey 1989). Let’s consider the general multivariate State Space system

$$\begin{cases} \mathbf{x}_t = \Phi \mathbf{x}_{t-1} + \mathbf{E} \mathbf{w}_t & \text{Transition Equations} \\ \mathbf{z}_t = \mathbf{H} \mathbf{x}_t + \mathbf{v}_t^* & \text{Observation Equations} \end{cases} \quad (1)$$

Here  $t$  is the time index measured in months;  $\mathbf{z}_t$  denotes the  $m$  dimensional output time series;  $\mathbf{x}_t$  is the  $n$  dimensional state vector;  $\mathbf{w}_t$  and  $\mathbf{v}_t^*$  are serial and mutually independent Gaussian noises of dimensions  $k$  and  $l$ , respectively, with constant covariance matrices; and  $\Phi$ ,  $\mathbf{E}$  and  $\mathbf{H}$  are known system matrices of appropriate dimensions. Without any loss of generality Eq. (1) is rewritten as

$$\begin{cases} \mathbf{x}_t = \Phi \mathbf{x}_{t-1} + \mathbf{E} \mathbf{w}_t \\ \begin{bmatrix} r_t \\ e_t \\ \mathbf{u}_t \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1 \\ \mathbf{H}_2 \\ \mathbf{H}^u \end{bmatrix} \mathbf{x}_t + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ \mathbf{v}_t \end{bmatrix} \end{cases} \tag{2}$$

Equations (1) and (2) are equivalent; the only difference is that the vector of output variables is separated in three parts:  $r_t$  and  $e_t$ , that are scalar variables, and  $\mathbf{u}_t$  that is a vector comprising the rest of  $m - 2$  variables. Matrices  $\mathbf{H}$  and  $\mathbf{v}_t^*$  have been split accordingly. In particular  $\mathbf{H}_1$ ,  $\mathbf{H}_2$  and  $\mathbf{H}^u$  stand for the first, second and the rest of rows of  $\mathbf{H}$  matrix in (1). Similarly  $v_{1,t}$ ,  $v_{2,t}$  and  $\mathbf{v}_t$  stand for the first, second and the rest of observed noises. Still, models (1) and (2) are exactly equivalent to a model in which the state vector is extended to include the first two output variables and the vector of transition noises is also extended with the first two observed noises

$$\begin{cases} \begin{bmatrix} r_t \\ e_t \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} 0 & 0 & \mathbf{H}_1 \Phi \\ 0 & 0 & \mathbf{H}_2 \Phi \\ \mathbf{0} & \mathbf{0} & \Phi \end{bmatrix} \begin{bmatrix} r_{t-1} \\ e_{t-1} \\ \mathbf{x}_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & \mathbf{H}_1 \mathbf{E} \\ 0 & 1 & \mathbf{H}_2 \mathbf{E} \\ \mathbf{0} & \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ \mathbf{w}_t \end{bmatrix} \\ \begin{bmatrix} r_t \\ e_t \\ \mathbf{u}_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}^u \end{bmatrix} \begin{bmatrix} r_t \\ e_t \\ \mathbf{x}_t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{v}_t \end{cases} \tag{3}$$

Time aggregation now may be incorporated by introducing a cumulator variable in system (3). This is a variable defined as

$$C_t = \begin{cases} 0, & t = \text{every january} \\ 1, & \text{otherwise} \end{cases}$$

The final model including the temporal aggregation constraints is

$$\begin{cases} \begin{bmatrix} r_t^A \\ e_t^A \\ \mathbf{x}_t \end{bmatrix} = \begin{bmatrix} C_t & 0 & \mathbf{H}_1 \Phi \\ 0 & C_t & \mathbf{H}_2 \Phi \\ \mathbf{0} & \mathbf{0} & \Phi \end{bmatrix} \begin{bmatrix} r_{t-1}^A \\ e_{t-1}^A \\ \mathbf{x}_{t-1} \end{bmatrix} + \begin{bmatrix} 1 & 0 & \mathbf{H}_1 \mathbf{E} \\ 0 & 1 & \mathbf{H}_2 \mathbf{E} \\ \mathbf{0} & \mathbf{0} & \mathbf{E} \end{bmatrix} \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ \mathbf{w}_t \end{bmatrix} \\ \begin{bmatrix} r_t^A \\ e_t^A \\ \mathbf{u}_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{H}^u \end{bmatrix} \begin{bmatrix} r_t^A \\ e_t^A \\ \mathbf{x}_t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{I} \end{bmatrix} \mathbf{v}_t \end{cases} \tag{4}$$

Variables  $\mathbf{u}_t$  in Eq. (4) are measured at a monthly rate, while  $r_t^A$  and  $e_t^A$  are only measured at the end of a year, i.e. they are observed every December, while they are missing the rest of the months. Thus, the value of the variable for one year is the sum

of the disaggregated measure for the 12 respective months. Comparison of Eqs. (3) and (4) clearly shows that  $r_t^A$  in (4) for any year is the accumulation along that year of the hypothetical monthly observations  $r_t$  in (3), if it were known. Every January the cumulator variable reset the monthly accumulation.

Beware that though models (3) and (4) look analytically similar, they have different time properties, since model (4) has one time varying system matrix due to the introduction of the cumulator variable. Exogenous variables may be added to the model by including a standard linear term to the observation equations in (4).

### 3.2 Models for the subsectors of the General Government

The basic model is of the Unobserved Component Model class known as the Basic Structural Model (Harvey 1989), that decomposes a set of time series in unobserved though meaningful components from an economic point of view (mainly trend, seasonal and irregular). The model is multivariate, and may be written as

$$\mathbf{z}_t = \mathbf{T}_t + \mathbf{S}_t + \mathbf{v}_t^* \tag{5}$$

Here  $\mathbf{T}_t$ ,  $\mathbf{S}_t$  and  $\mathbf{v}_t^*$  denote a trend, seasonal and irregular components, respectively. Equation (5) is in fact a set of observation equations in a State Space system, which has to be completed by the standard transition or state equations. 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, relations are only allowed through the covariance structure of the vector noises  $\mathbf{w}_t$  and  $\mathbf{v}_t^*$ , 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 state equations qualify the dynamic behaviour of the components, and a full model may be built by block concatenation of the individual components. The transition equations for models of the trend and seasonal components are a Local Linear Trend and the Trigonometric Seasonal in Eq. (6),

$$\left\{ \begin{array}{l} \text{Trend :} \quad \begin{pmatrix} \mathbf{T} \\ \mathbf{D} \end{pmatrix}_t = \Phi_T \begin{pmatrix} \mathbf{T} \\ \mathbf{D} \end{pmatrix}_{t-1} + \mathbf{E}_T \begin{pmatrix} \mathbf{w}_0 \\ \mathbf{w}'_0 \end{pmatrix}_t \\ \text{Seasonal :} \quad \mathbf{S}_t = \sum_{i=1}^6 \mathbf{S}_{it}; \quad i = 1, 2, \dots, 6 \\ \quad \quad \quad \begin{pmatrix} \mathbf{S}_i \\ \mathbf{S}'_i \end{pmatrix}_t = \Phi_i \begin{pmatrix} \mathbf{S}_i \\ \mathbf{S}'_i \end{pmatrix}_{t-1} + \mathbf{E}_T \begin{pmatrix} \mathbf{w}_i \\ \mathbf{w}'_i \end{pmatrix}_t, \quad \omega_i = \frac{2\pi i}{12} \end{array} \right. \tag{6}$$

with

$$\Phi_T = \begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}; \mathbf{E}_T = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}; \Phi_i = \begin{pmatrix} \cos \omega_i \mathbf{I} & \sin \omega_i \mathbf{I} \\ -\sin \omega_i \mathbf{I} & \cos \omega_i \mathbf{I} \end{pmatrix}, \quad i = 1, 2, \dots, 6$$

In Eq. (6)  $\mathbf{D}_t$  and  $\mathbf{S}'_{it}$  are additional states necessary to define the components;  $\mathbf{I}$  and  $\mathbf{0}$  are the identity matrix and a square block of zeros of dimension  $m$ ;  $\mathbf{w}_j$  and  $\mathbf{w}'_j$  ( $j = 0, 1, \dots, 6$ ) are multivariate Gaussian white noises serially independent and independent of each other; and  $\omega_i$  ( $i = 1, 2, \dots, 6$ ) are the fundamental frequency of the seasonal component and its harmonics.

The State Space form of a full BSM model written as Eq. (4) consists of a state vector  $\mathbf{x}_t$  composed by the concatenation of all the states regarding the trend components ( $\mathbf{T}_t$  and  $\mathbf{D}_t$ ) and all the states regarding seasonal components  $\mathbf{S}_{it}$  and  $\mathbf{S}'_{it}$  ( $i = 1, 2, \dots, 6$ ) in Eq. (6). Matrix  $\Phi$  is the block concatenation of matrices  $\Phi_T$  and  $\Phi_i$ ,  $i = 1, 2, \dots, 6$ ;  $\mathbf{E}$  is the block concatenation of seven  $\mathbf{E}_T$  matrices; and  $\mathbf{H}$  is the horizontal concatenation of seven  $[\mathbf{I} \ \mathbf{0}]$  matrices.

Within this general structure, the specific models estimated for each one of the subsectors of the General Government will be:

*Model CGBlock*: model of type (6) for the CG sector.  $r_t$  and  $e_t$  refers to annual total revenue and total expenditure of the CG sector in ESA95 terms.  $\mathbf{u}_t$  comprises eight monthly variables: four on the revenue side (State sector direct taxes, VAT, other indirect taxes, and other revenues) and four on the expenditure side (State sector government consumption, interest payments, capital expenditure, and other expenditures).

*Model SSblock*: model of type (6) for the SS sector.  $r_t$  and  $e_t$  refers to annual total revenue and total expenditure of the Social Security sector in ESA95 terms.  $\mathbf{u}_t$  comprises one monthly revenue variable and one monthly expenditure variable.

*Model LGBlock*: model of type (6) for the LG sector.  $r_t$  and  $e_t$  refers to annual total revenue and total expenditure of the sector in ESA95 terms.  $\mathbf{u}_t$  only contain one monthly revenue indicator.

In order to reduce the dimensionality of the models (in terms of parameters to be estimated), some constraints have been imposed among the unobserved components of  $r$ ,  $e$ , and  $\mathbf{u}$ , that can be partly rationalized on the basis of economic priors. First, the structure of the covariance matrix for the noises affecting the seasonal components is assumed diagonal, i.e. all the seasonal components are assumed independent of each other. The main motivation for this assumption is the fact that cash and accrual variables differ in the way they are computed within the year, i.e. seasonal components reflect mainly differences in accounting standards.

Second, the covariance matrices for the noises in the trend levels, the trend slopes and the irregulars are specified according to the following constraints: (1) the ones corresponding to  $r$  and  $e$  are correlated, (ii) the block of revenue indicators in  $\mathbf{u}$  is modelled jointly but independent of the block of expenditures, (3) the block of expenditure indicators in  $\mathbf{u}$  is modelled jointly but independent of revenues, (4)  $r$  is only related to the corresponding block of revenue indicators, (5)  $e$  is only related to the corresponding block of expenditure indicators. Thus, we allow for full interdependence of total revenues and total expenditures in ESA terms (the key variables), while at the same time incorporate dimensionality-reducing constraints between the blocks of revenue and expenditure indicators.

### 3.3 Model including the General Government and all the subsectors

A final model (*Model GG*) will be a joint model for the full dataset, with the following properties: (1) quarterly information about GG total revenue and total expenditure from the first quarter of 1995 onwards is incorporated into the model, (2) it is built by orthogonal block concatenation of the individual models for the subsectors discussed in the previous subsection.

*Model GG* is specified in order to produce a joint explanation of the problem. The model is multivariate in a weak sense, since the only connection among different sectors is given by the fact that the sum of the three subsectors' total revenue and expenditure is equal to the total revenue and expenditure of the GG sector.<sup>4</sup> In addition to the joint estimation of the blocks, the joint model allows the integration of the available quarterly information for the General Government sector while, as stated in a previous section, there is no quarterly information available for the subsectors. Thus, a new cumulator variable has to be defined

$$C_t^* = \begin{cases} 0, & t = \text{every January (monthly data) or} \\ & \text{first month of every quarter (quarterly data)} \\ 1, & \text{otherwise} \end{cases}$$

In order to explain how *Model GG* is set up, consider a system formed by the concatenation of three systems of type (6) for the CG, SS and LG sectors (models *CGBlock*, *SSBlock* and *LGBlock*, the latter with  $\mathbf{u}_t$  empty) in Eq. (7)

$$\begin{cases} \mathbf{X}_t = \bar{\Phi}_t \mathbf{X}_{t-1} + \bar{\mathbf{E}} \mathbf{W}_t \\ \mathbf{Z}_t = \bar{\mathbf{H}} \mathbf{X}_t + \bar{\mathbf{C}} \mathbf{V}_t \end{cases} \tag{7}$$

Here capital letters stand for the vertical concatenation of the state vector, noises and endogenous variables, while bar letters indicate diagonal block concatenation of corresponding matrices. General Government ESA95 variables have to be added to the system, bearing in mind that the data is annual up to 1994 and quarterly afterwards. Application of time aggregation techniques by means of the cumulator variable  $C_t^*$  gives

$$\begin{cases} \begin{bmatrix} R_t \\ E_t \\ \mathbf{X}_t \end{bmatrix} = \begin{bmatrix} C_t^* & 0 & \bar{\mathbf{H}}_R \bar{\Phi}_t \\ 0 & C_t^* & \bar{\mathbf{H}}_E \bar{\Phi}_t \\ \mathbf{0} & \mathbf{0} & \bar{\Phi}_t \end{bmatrix} \begin{bmatrix} R_{t-1} \\ E_{t-1} \\ \mathbf{X}_{t-1} \end{bmatrix} + \begin{bmatrix} \bar{\mathbf{H}}_R \bar{\mathbf{E}} \\ \bar{\mathbf{H}}_E \bar{\mathbf{E}} \\ \bar{\mathbf{E}} \end{bmatrix} \mathbf{W}_t \\ \begin{bmatrix} R_t \\ E_t \\ \mathbf{Z}_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & \mathbf{0} \\ 0 & 1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \bar{\mathbf{H}} \end{bmatrix} \begin{bmatrix} R_t \\ E_t \\ \mathbf{X}_t \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \\ \bar{\mathbf{C}} \end{bmatrix} \mathbf{V}_t \end{cases} \tag{8}$$

<sup>4</sup> More complex models were entertained, in which relations among sectors were specified explicitly via the covariance matrices of the components, but the estimation problems at that stage were too discouraging. In the end, the information added for the estimation of this model is quarterly information for part of the sample in two aggregated time series, i.e. too little information for the amount of additional parameters to be estimated.



Here  $R_t$  is annual/quarterly GG revenue;  $E_t$  is GG expenditure;  $\bar{\mathbf{H}}_R$  is a row vector formed by the addition of rows in  $\bar{\mathbf{H}}$  corresponding to the aggregated revenue of CG, SS and LG sectors; and  $\bar{\mathbf{H}}_E$  is a row vector formed by the addition of rows in  $\bar{\mathbf{H}}$  corresponding to the aggregated expenditure of CG, SS and LG.

### 3.4 Estimation of State Space models

Given any model written in State Space form, the estimation problem consists of finding the first and second order moments (mean and variance) of the state vector, conditional to all the data in a sample. The widespread general tools to perform this operation in a State Space framework are the Kalman Filter and the Fixed Interval Smoothing algorithms (full accounts can be seen, e.g. in [Harvey 1989](#); [Pedregal and Young 2002](#)).

One classical problem is that the application of the recursive algorithms above requires the knowledge of all the system matrices. In most cases, though not always necessarily, there will be some unknown elements, usually called hyper-parameters, which must be estimated by efficient methods. There exist a number of ways to deal with this problem, though Maximum Likelihood in time domain is the most widespread estimation method, mainly because of its strong theoretical basis. Assuming that all the disturbances in the state space form are normally distributed, the Log-likelihood function can be computed using the Kalman Filter via ‘prediction error decomposition’ ([Schweppe 1965](#); [Harvey 1989](#)). There is not any need for especial considerations about the estimation of models with time aggregation constraints. Since this is a classical estimation procedure the topic is not pursued further in this paper (see [Harvey 1989](#); [Pedregal and Young 2002](#)). Some diagnostic checks of the annual revenue and expenditure residuals of the models estimated by Maximum Likelihood are shown in [Table 2](#). All tests show clearly absence of any kind of problem in the residuals, since there is no evidence of serial dependence, residuals are Gaussian and no evidence of heteroskedasticity can be found.

## 4 Empirical exercise

### 4.1 Forecast performance statistics: current year forecasts

In order to replicate the real-time constraints faced by real-time analysts, we adopt the timing rules displayed in [Table 3](#), following the standard dates of dissemination of data at the different frequencies. In the table we show the available information in each quarter of a given year. Annual ESA95 figures for year  $t - 1$  are first released by the national statistical office in March/April of year  $t$  but the validation processes by Eurostat of figures reported by national statistical agencies render April/May as the actual date in which usable/reliable figures are available to an outside analyst. Thus, from a quarterly observation perspective, it is fair to assume that the annual figure for year  $t - 1$  is only available in the second quarter of year  $t$ . In a related fashion, the quarterly ESA95 figure for the fourth quarter of year  $t - 1$  would only be available in the course of the second quarter of year  $t$ . Regarding monthly cash accounts, we follow

**Table 2** Diagnostic checks of annual revenue and expenditure residuals of all models

Sector	Variables	<i>Q</i> (1 year)	<i>Q</i> (2 years)	Jarque-Bera	<i>P</i> (Jarque-Bera)	<i>P</i> (Heterosk.)
Models for individual sectors (Model CG block, Model SS block, Model LG block)						
CG	Revenues	0.46	0.65	6.09	0.04	0.11
	Expenditures	0.14	1.59	1.22	0.54	0.45
SS	Revenues	0.15	1.59	4.07	0.13	0.45
	Expenditures	0.02	0.15	1.49	0.47	0.47
LG	Revenues	0.32	1.22	2.47	0.29	0.35
	Expenditures	0.05	0.67	0.65	0.72	0.25
Overall model (Model GG)						
CG	Revenues	0.87	1.06	1.05	0.59	0.24
	Expenditures	0.003	0.25	1.66	0.44	0.49
SS	Revenues	0.09	2.04	3.92	0.14	0.43
	Expenditures	0.14	1.28	1.15	0.56	0.48
LG	Revenues	1.56	1.70	0.84	0.65	0.32
	Expenditures	0.17	2.23	1.58	0.45	0.24

*Q* (1 year) and *Q* (2 years) stand for the Ljung and Box *Q* statistic of serial correlation for 1 and 2 years, respectively. Jarque-Bera and *P* (Jarque-Bera) are a gaussianity test and its *P* value, respectively. *P* (Heterosk) is the *P* value of a standard ratio of variances test

**Table 3** Timing rules

	Q1 year <i>t</i> (March)	Q2 years <i>t</i> (June)	Q3 year <i>t</i> (September)	Q4 year <i>t</i> (December)
Available annual (A)	$A_{t-2}$ (March)	$A_{t-1}$ (April)	$A_{t-1}$	$A_{t-1}$
Available quarterly (Q)	$Q_{3,t-1}$	$Q_{4,t-1}$ (April)	$Q_{1,t}$	$Q_{2,t}$
Available monthly (M)	Jan. <i>t</i>	Jan.–April <i>t</i>	Jan.–July <i>t</i>	Jan.–October <i>t</i>

the assumption of availability with a lag of two months. We deem this convention also as a fair heuristic representation of reality.

In order to carry out the forecasting exercise, we add to the list of models described in the previous section the following models: (1) annual random walk (ARW henceforth), (2) quarterly random walk (QRW henceforth), consisting of adding the last four quarters available of the series when quarterly data is available (i.e. General Government) and equals the ARW forecasts when the forecast origin coincides with the end of each year.<sup>5</sup> The QRW alternative allows to test against an alternative

<sup>5</sup> Random walk models were selected because they are standard, benchmark alternatives in the forecasting literature. In the case of our particular empirical setup other standard benchmark alternatives, like a AR(1) applied to quarterly ESA95 series, were discarded on the basis that, due to the limited sample available (for quarterly data the sample starts in 1995Q1) the estimation of simple AR(1) models turned out to show high RMSEs. This finding is consistent with the fact that, given the non-stationary nature of the modelled series (series in nominal terms), more complex ARIMA structures could have been needed. For the sake of simplicity and focus of the empirical exercises, we preferred to stick to standard, simpler alternatives like the QRW.

purely based on quarterly GG information. The ARW alternative is the standard naive benchmark.

Then 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 2001 to September 2007 (this makes up to 28 projections at each forecast horizon). The first forecast origin is March 2001, and following the timing convention outlined before (see Table 3) the annual information available covers up to the year 1999, the quarterly information up to 2000:Q3, and the monthly information up to January 2001. The second forecast origin is June 2001, with annual information up to 2000, quarterly up to 2000:Q4 and monthly up to January–April 2001. Then we move the forecast origin to September 2001 and so on and so forth until September 2007. We leave out of the exercises in this section the year 2008, that will be analysed in isolation in a subsequent section.

Finally, we present two standard, quantitative measures of forecasting performance. Firstly, we look at the ratio of the root mean squared errors (RMSE) of the different alternative models with respect to the ARW model. Secondly, we also look at the Diebold and Mariano test, and test for the null hypothesis of no difference in the accuracy of two competing forecasts. We make sure that a reasonable proportion of the sample is employed when the first out-of-sample forecast is computed to reduce the bias generated by ignoring parameter uncertainty (see for example Clark and McCracken 2001); the forecasting exercise is performed on the moving window 2001–2008, while the full sample covers 1985–2008.

We focus on the forecast performance of annual projections, i.e. forecasts generated from each forecast origin for the end of the current year. Table 4 shows the results for the end-of-the-year forecasts performed using the alternative models. It shows first the RMSE ratios of the different alternative models to the ARW alternative and, second the Diebold–Mariano tests of equal forecast accuracy for each pair of models.

One set of evidence on the usefulness of intra-annual information is provided by the fact that the RMSE ratios of model QRW over ARW is 0.85, implying that a blind use of quarterly ESA95 information would improve forecasts for the end of the year. As regards the rest of RMSE ratios the results clearly show that all alternatives with intra-annual update beat the ARW. The ratios range from 0.32 of *Model GG* to the 0.85 of *LGblock*. The best results are those for the overall GG, due to the fact that the best forecasts by sectors correspond to CG, that represent the biggest proportion of the overall General Government sector. In 2006 the proportion of revenue (expenditure) was 43.8% (43.7%) for CG, 37.5% (36.2%) for SS and 18.7% (20.1%) for LG. As regards the ratios between pairs of the other alternatives except ARW, all models beat QRW. The RMSE of Model GG and the addition QRW are 0.37 and 0.43, respectively. Regarding the results of the Diebold–Mariano test, all models outperform the ARW model significantly. Either model GG or the sum of forecasts of the three individual models for the subsectors produce significant better forecasts than the QRW model. It is also worth noticing that model GG systematically produces better RMSE ratios than the block models, and thus we may conclude that including ESA95 quarterly information into the analysis is important for improving forecasting results.

**Table 4** Forecast performance statistics: current year forecasts

	RMSE/ARW	Diebold–Mariano tests		CG + SS + LG Block
		ARW	Full	
<b>GG</b>				
Full	0.318	-2.308**	–	–
CG + SS + LG Block	0.366	-1.988**	0.344	–
QRW	0.851	-1.667*	2.403**	2.191**
<b>CG</b>				
Full	0.439	-2.139**	–	–
CG Block	0.553	-1.685*	0.176	–
<b>SS</b>				
Full	0.597	-1.988**	–	–
SS Block	0.663	-2.292**	0.192	–
<b>LG</b>				
Full	0.838	-1.793*	–	–
LG Block	0.853	-1.678*	0.646	–

Ratio of RMSE of each method against an annual random walk (ARW), and Diebold–Mariano test for the null hypothesis of equal forecast accuracy of two forecast methods. The numbers in each cell of the DM test represents the loss differential of the method in its vertical column as compared to the method in the horizontal line

The Diebold–Mariano statistic follows a  $N(0,1)$  distribution. A single (double) asterisk denotes rejection of the null hypothesis at the 10% (5%) significance level

The main lessons to be drawn from Table 4 are: (1) all models with intra-annual update beat ARW, (2) models including monthly indicators beat QRW, (3) overall, model GG tends to give better results for the fiscal variables of the individual subsectors than the block model counterparts.

#### 4.2 Forecast performance statistics: now-casting and current quarter forecasts

In Table 3 we focus on forecasts for the whole year, while in Table 5 we turn to the related issue of now-casting and one-quarter-ahead forecasts, also relevant for decision-taking in real time.

In real-time, policy makers face the following problems. First, from a public finance point of view, annual targets are set in terms of National Accounts’ definitions and refer to the General Government sector, while the flow of incoming information tend to follow cash recording practices (the case of monthly Social Security and Regional governments’ figures) or is related to a given sub-sector of the General Government (the case of ESA95 monthly Central Government figures). At the same time, from the point of view of macroeconomic forecasting, quarterly national accounts form the articulating framework of short-term macroeconomic forecasts.

From these perspectives, the methodology and the approach developed in this paper would provide a useful tool for policy makers and real-time forecasters if, on the basis of incoming fiscal information, our models were able to provide fair current quarter

**Table 5** Forecast performance statistics: nowcasting and current quarter forecasts

	RMSE/QRW	Diebold–Mariano tests	<i>P</i> value
Nowcasts	0.565	−3.495**	0.0005
One quarter ahead	0.727	−3.074**	0.0021

Ratio of RMSE measures for quarterly General Government sector of model GG (Full) (nowcasts and one quarter ahead) against a quarterly random walk (QRW) forecast, and Diebold–Mariano test for the null hypothesis of equal forecast accuracy of two forecast methods. Nowcasts are computed by assuming all the monthly information up to the quarter to be forecast; one quarter ahead forecasts are computed by assuming the information available at that point in time according to Table 2. The numbers in each cell of the DM test represents the loss differential of the method in its vertical column as compared to the method in the horizontal line

The Diebold–Mariano statistic follows a  $N(0,1)$  distribution. A single (double) asterisk denotes rejection of the null hypothesis at the 10% (5%) significance level

estimates (now-casting if all the months of the quarter were available) and fair one-quarter-ahead projections of fiscal aggregates in National Accounts.

These are exactly the exercises whose results are provided in Table 5. Now-casts are computed by assuming all the monthly information up to the quarter to be forecast; one quarter ahead forecasts are computed by assuming the information available at that point in time according to Table 3; QRW consists of replacing the forecast for the next quarter by the latest quarter of the same type available within the dataset. The ratio of RMSEs is the ratio of the RMSE of model GG (now-casts and one quarter ahead) against a quarterly random walk (QRW) forecast, and the Diebold–Mariano test shows a comparison of the same two methods. The numbers clearly show how model GG outperforms a quarterly random walk in the two investigated respects (now-casting and one-quarter ahead forecasts); the relative behaviour of model GG is even better than the one displayed for current year forecasts (see Table 4).

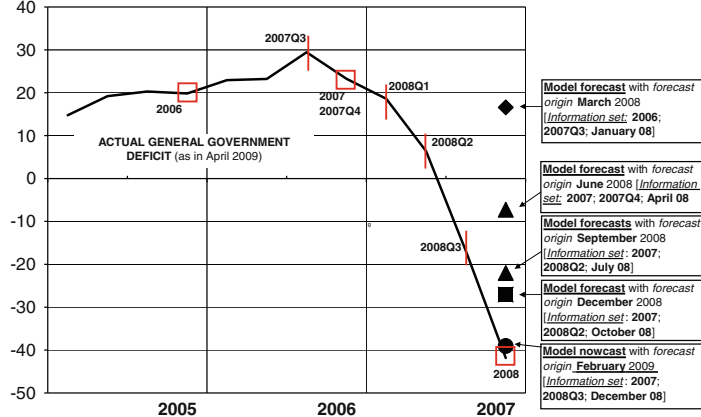
#### 4.3 A real-time illustration

Figure 5 completes the information shown in Tables 4 and 5 by presenting the behaviour of the models in the most recent period, the year 2008, that was not included in the previous forecasting exercises. We decided to leave 2008 out of the forecast sample because it provides an episode of a dramatic deterioration of the general government balance that moved from a surplus of 1.9% of GDP in 2008 to a deficit of 4.1% of GDP in 2008, according to the revised estimates published in Autumn 2009.

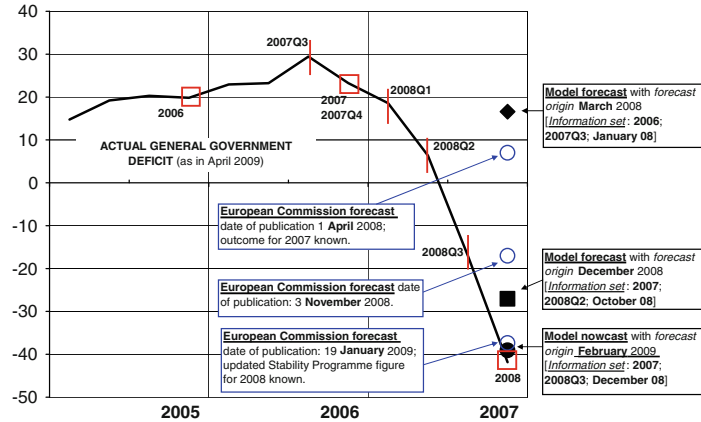
We illustrate the mechanical behaviour of our time series models, against published government targets/estimates, and published forecasts by the European Commission (EC). This implies leaning against the mixed-frequency models, as in the course of the year a number of discretionary fiscal policy measures affecting 2008 were announced and implemented. In real-time, these measures, forward-looking in nature, could have been included in our time series models. The exercise is a full real-time exercise, in that the series/forecasts used in all cases were those available in real-time.

In the first quarter of 2008 the government still presented a general government surplus as a target; with information up to February/March of 2008, the models would have

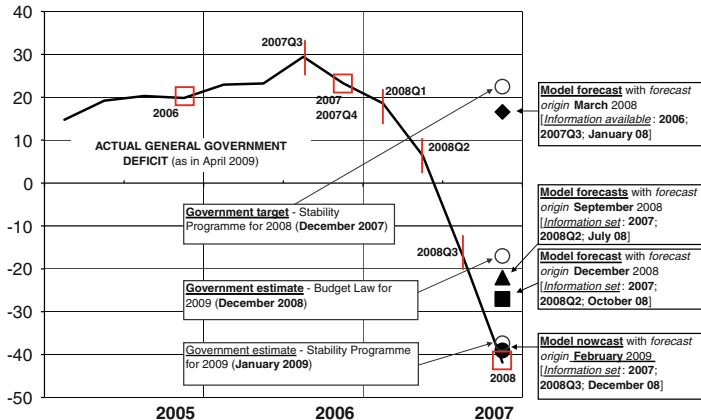
**Panel A**



**Panel B**



**Panel C**



**Fig. 5** Spanish General Government balance (billion euros) and alternative sets of projections for the end of 2008: Full model projections, European Commission projections and Government projections. Forecast origin and information set available shown

provided little signals of the appearance of a deficit in 2008, even though the estimates were below the government's (see Panel C). With subsequent information covering up until April, and disregarding discretionary fiscal policy measures announced at that time, the time series models would have already pointed to a small deficit for 2008 (triangle in Panel A). Indeed, EC projections, taking account of policy measures announced in April, were already capturing the turning point, even though by April 2008 the EC still projected a small surplus. For the subsequent months, the mechanical application of the time series methods would have traced very well the deterioration subsequently incorporated by EC projections and government successive estimates for 2008.

With all the possible available information at the monthly frequency included in the model (December 2008 for the monthly indicators), the model would have now-casted a deficit quite close to the final (first) estimate provided by the statistical agency (Panel A of Fig. 5, "Model now-cast" box). This is the case, notwithstanding the fact that intra-annual fiscal information for the Regional and Local governments is quite scarce in our models (as in reality), and a good part of the deterioration of the deficit of the general government, ex-post, was due to the more-than-usual negative contribution of these sub-sectors. It is interesting to note that the way intra-annual developments in the Regional and Local governments are internalised in the model is via quarterly General Government figures. Taking into account that, given our tight timing rules, only 2008Q3 was assumed to be available to produce the "Model now-cast" projections for 2008, one can conclude that the estimates of our model are reasonable enough, and promising for real-time decision taking.

## 5 Conclusions

We construct multivariate, state-space mixed-frequencies models for the Spanish General Government sector made up of blocks for each one of its subsectors: the Central Government sector, the Social Security Sector and the aggregate of the Regional and Local government sectors. Each block is modelled through its total revenue and expenditure categories, and encompasses a number of indicators for each subsector, depending on data availability. In this sense, our model makes a fair coverage of the Central Government sector, while the blocks for the Social Security sector and the Regional and Local governments encompass a more limited coverage due to data availability. The mixed-frequencies approach is particularly relevant for the case of Spain, given the specific data coverage of the different subsectors of the general government.

All in all, we provide evidence that our models are appropriate tools: (1) for real-time monitoring of fiscal policies with a focus on quarterly developments of general government figures, (2) for the monitoring of general government sub-sectors for which intra-annual data coverage is limited (Regional and Local governments), (3) to bridge (translate) into National Accounts available monthly information for the subsectors of the general government.

**Acknowledgments** We thank seminar participants at the Bank of Spain, the Editor (Eduardo Ley) and two anonymous referees for useful comments and suggestions. Pedregal acknowledges financial support of the Spanish Education and Science Ministry (project: SEJ2006-14732 (ECON)).

## 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 Papers 06/116, International Monetary Fund
- Chow GC, Lin A (1971) Best linear interpolation, distribution and extrapolation of time series by related series. *Rev Econ Stat* 53:372–375
- Clark TE, McCracken MW (2001) Tests of equal forecast accuracy and encompassing for nested models. *J Econom* 105:85–110
- Clements MP, Galvão AB (2008) Macroeconomic forecasting with mixed frequency data: Forecasting US output growth. Working Paper 616, Department of Economics, University of London
- European Commission (2008) Economic forecast—Autumn 2008. *European Economy* 6/2008, Directorate-General for Economic and Financial Affairs
- Frale C, Veredas D (2008) A monthly volatility index for the US real Economy. *ECORE DP* 2008/15
- Ghysels E, Santa-Clara P, Valkanov R (2006) Predicting volatility: getting the most out of return data sampled at different frequencies. *J Econom* 131:59–95
- Gordo L, Hernández de Cos P (2001) The Financing Arrangements for the Regional (Autonomous) Governments for the Period 1997–2001. Working Paper 0003, Research Department, Bank of Spain
- Guerrero VM (2003) Monthly disaggregation of a quarterly time series and forecasts of its unobservable monthly values. *J Official Stat* 19:215–235
- Harvey A (1989) *Forecasting structural time series models and the Kalman Filter*. Cambridge University Press, London
- Harvey AC, Chung CH (2000) Estimating the underlying change in unemployment in the UK. *J R Stat Soc Ser A* 163:303–339
- Hyung N, Granger C (2008) Linking series generated at different frequencies. *J Forecast* 27:95–108
- Jonung L, Larch M (2006) Fiscal policy in the EU: are official output forecasts biased? *Econ Policy* 491–534
- Leal T, Pérez JJ, Tujula M, Vidal JP (2008) Fiscal forecasting: lessons from the literature and challenges. *Fiscal Stud* 29:347–386
- Leal T, Pérez JJ (2009) Un sistema ARIMA con Agregación Temporal para la Previsión y el Seguimiento del Déficit del Estado. *Hacienda Pública Española* 190:27–58
- Liu H, Hall SG (2001) Creating high-frequency national accounts with state space modelling: a Monte Carlo experiment. *J Forecast* 20:441–449
- Mariano RS, Murasawa Y (2003) A new factor of business cycles based on monthly and quarterly series. *J Appl Econom* 18:427–443
- Moauero F, Savio G (2005) Temporal disaggregation using multivariate structural time series models. *Econom J* 8:214–234
- Moulin L, Wierds P (2006) How credible are multiannual budgetary plans in the EU? *Fiscal Indicators*, Banca d'Italia, pp 983–1005
- Onorante L, Pedregal DJ, Pérez JJ, Signorini S (2009) The usefulness of infra-annual government cash budgetary data for fiscal forecasting in the euro area. *J Policy Model* (in press)
- Pedregal DJ, Pérez JJ (2009) Should quarterly government finance statistics be used for fiscal surveillance in Europe? *Int J Forecast* (in press)
- Pedregal DJ, Young PC (2002) Statistical approaches to modelling and forecasting time series. In: Clements M, Hendry D (eds) *Companion to economic forecasting*. Blackwell Publishers, Oxford, pp 69–104
- Pérez JJ (2007) Leading indicators for euro area government deficits. *Int J Forecast* 23:259–275
- Pina A, Venes N (2008) The political economy of EDP fiscal forecasts: an empirical assessment, Working Papers 2008/23, Department of Economics at the School of Economics and Management (ISEG), Technical University of Lisbon
- Proietti T, Moauero F (2006) Dynamic factor analysis with nonlinear temporal aggregation constraints. *J R Stat Soc Ser C Appl Stat* 55:281–300
- Schweppe F (1965) Evaluation of likelihood function for Gaussian signals. *IEEE Trans Info Theory* 11:61–70
- Silvestrini A, Salto M, Moulin L, Veredas D (2008) Monitoring and forecasting annual public finance deficit every month: the case of France. *Empir Econ* 34(3):493–524
- Strauch R, Hallerberg M, Von Hagen J (2004) Budgetary Forecasts in Europe—the track record of stability and convergence programmes. European Central Bank WP 307