T E C H N I C A L

R E P O R T

0469

USING INTRA ANNUAL INFORMATION TO FORECAST THE ANNUAL STATE DEFICITS. THE CASE OF FRANCE

MOULIN, L., SALTO, M., SILVESTRINI, A. and D. VEREDAS

I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

CORE DISCUSSION PAPER
2004/48

USING INTRA ANNUAL INFORMATION TO FORECAST THE ANNUAL STATE DEFICITS. THE CASE OF FRANCE

Laurent Moulin[1], Matteo Salto[2], Andrea Silvestrini[3] and David Veredas[4]

July 2004

Abstract

We develop a methodology for using intra-annual data to forecast annual budget deficits. Our approach aims at improving the accuracy of the deficit forecasts, a relevant issue to policy makers in the Eurozone and at proposing a replicable methodology using at best public quantitative information on budgetary data. Using French data on government (State) revenues and expenditures, we estimate intra-annual monthly ARIMA models for all the items of the central government revenues and expenditures. Next, applying temporal aggregation techniques, we infer parameters of the annual models from the estimated parameters of the intra-annual models. These parameters incorporate all the intra-annual information. Finally, we do one period ahead predictions. We are able to update the annual deficit forecast as soon as new monthly data are available. This allows us to detect possible slippages in central government finances.

Keywords: French State deficit, temporal aggregation, intra-annual, forecasting.
JEL classification: C22, C53, E62, H60

[1]Directorate-General ECFIN, European Commission.
[2]Directorate-General COMP, European Commission.
[3]Department of Statistical Sciences, University of Perugia, Italy.
[4]CentER, Department of Econometrics and Operations Research, Tilburg University, The Netherlands, and CORE, Université catholique de Louvain, Belgium.
Corresponding address: David Veredas, Tilburg University, Dept. Econometrics & OR. P.O. Box 90153. 5000LE Tilburg, The Netherlands. Phone: +31 13 466 3151. Fax: +31 13 466 3280. Email: d.veredas@uvt.nl
The authors wish to thank Juan Toro, who provided the data. While remaining responsible for any error in this paper, the authors are grateful to Helena Beltran-Lopez, Javier Perez and Juan Toro (data provider), as well as to the seminar audiences at Tilburg University, European Commission, catholic University of Louvain, CentrA workshop on Public Finance (Sevilla, 2003) and congress of Forecasting Financial Markets (Lodz, 2004) for useful remarks and suggestions on previous versions. This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction

The strengthening of multi-lateral budgetary surveillance that followed the adoption of the euro has increased the need for timely and accurate forecasts of budgetary developments in the EU. Indeed, the legal texts (European Union Treaty, Stability and Growth Pact) impose limits to the yearly budget deficits of the public sector. The surveillance mechanism requires notably that, when the general government deficit of a member state risks to exceed 3% of GDP, the European Council - that is the body in charge of the coordination of budgetary policies - may take policy decisions and ask the member state to act in order to avoid that general government deficit becomes "excessive".

There are two situations in which the European Council needs to base its decisions on estimates for the government finances developments rather than on actual figures: when a decision has to be taken for the years ahead, and when a decision has to be taken for the ongoing year. In the latter case a practical problem arises, that no intra-annual statistics for the general government deficit are available. And that no reliable estimate of the general government deficit of the ongoing year is available before March of the next year. This leads the Council to base its decisions on forecasts, which obviously should have some desirable properties of which reliability, transparency, duplicability and ease of replicability with, are the most relevant.

In the usual approach used by governments and international organizations - which will be named the "traditional" approach - government finances forecasts are based on the estimate of three types of parameters: i) the macroeconomic developments, ii) the fluctuations in the elasticities of taxes to the tax bases and iii) the developments in government expenditures.

In order to predict tax revenues, forecasters usually express them as the product of a tax base times an 'apparent' tax rate. Tax bases are in general proxied by a function of macroeconomic variables - usually predicted using a large macroeconomic model. Fluctuations in the effective tax rates - which are often reinterpretated as variation of the elasticities of tax revenues to the tax bases, a simple algebric transformation - depend on several parameters: the presence of "discretionary measures" (like changes in the legislation which increase / decrease taxes), the composition of growth, specific aspects of the tax system (leads and lags in tax revenues, tax ceilings, tax progression), and other parameters (private agents' expectations for example). In order to forecast developments in tax rates, the independent forecaster uses educated guesses based on government's estimates, its knowledge of the legislation, and conventional assumptions; experience shows that tax rates fluctuations are one of the main sources of error when forecasting government finances developments. Finally, projecting government expenditures is probably the most difficult part. Here, the traditional forecaster relies on a mixture of educated guess based on the information provided by the government, simple rules based on past respect/non respect of budgetary targets by the government and estimation of the evolution of certain entitlements (pensions, unemployment benefits for examples), which depend on macroeconomic, legislative and demographic factors. This technique reflects the fact that the largest part of expenditures are under the direct control of the authorities, the behaviour of which is difficult to predict.

The traditional approach has one undisputable advantage: forecasters can present simply their results and explain them to policy makers, because they can link clearly government finances, projected developments to the evolution of macroeconomic variables (i.e. developments "independent" from fiscal policy choices) and discretionary measures (i.e. to developments "depending "on new fiscal policy decisions). This, together with the fact that it relies on a number of intuitive and simple assumptions, makes this method very convenient, given that, after all, these forecasts are supposed to help policy makers in taking decisions. The second advantage of this method is that it potentially takes into account all kind of information known at the moment of the forecast and can therefore show the impact of the most recent news on the predictions. Finally, this method is fully coherent with national accounting which is used for producing the macroeconomic figures and which is used in the legal basis of the framework defining the coordination of fiscal policy in the EU.

However, this method has also various drawbacks. First, it often leads to poor results, partly due to the fact that the results depend on a number of "conventional" assumptions and partly due to the large errors generated by the overlapping of the numerous steps of the forecast, each of them involving various assumptions. Second, the forecast procedure is complex and requires many resources, both human and informational resources, which makes it not easily replicable in the course of the year. Third, the role played by the existing figures at a frequency which is less than yearly is not clear when intra-annual forecasts are made. Fourth, the traditional approach leaves a large room for judgemental choices which can easily be not transparent and subject to political pressures. Finally, in this approach it is not clear what is the accuracy of the forecast and the properties of the resulting estimator cannot be defined.

In this article, we propose to apply temporal aggregation techniques in forecasting within-year budgetary developments to best exploit the existing intra-annual information. We exploit the existence of monthly budgetary statistics - the cash expenditures and revenues components for the State deficit - in France, our example country, which give useful indications about the situation of the government deficit of the ongoing year. This method allows to improve upon certain drawbacks of the traditional approach and of the traditional alternative time series approach. Moreover, we can show that the forecast based on our methodology is promising in that, by the beginning of the second quarter of the year, it allows to make predictions for the public deficit of the current year that are reliable.

The idea of applying purely statistical methods to existing intra-annual data to forecast the annual government deficits is already present in the literature, in particular in the USA, where the existence of deficit ceilings is a common feature of the majority of states. Examples are the papers by Bretschneider et al. (1989), Nazmi and Leuthold (1988), Fullerton (1989), Grizzle and Klay (1994) and Lawrence et al. (1998). They compare the accuracy of the forecasts made using time series techniques with the forecasts made using the traditional methodology. In Europe, given the more recent introduction of deficit rules and the scarcity of intra-annual budgetary data, few studies exist that exploit existing intra-annual information. Exceptions are Kinnunen (1999), who uses existing Finnish data, Camba-Mendez and Lamo (2002) and Perez (2003), with monthly cash figures for the State deficit. These papers share the approach of "letting the data speak", in the sense that no place is left for judgemental analysis, but the estimates are made using time series analysis with no other information except the own past of the variables.

Our paper is based on the same logic. However, it improves upon the existing literature by using for the first time temporal aggregation techniques, which seems to be a natural choice when data of a higher frequency (monthly) are used to predict a lower frequency (annual) variable. In practice, we estimate linear time series models, i.e. ARIMA with outliers detection, for each revenue and expenditure component at monthly frequency. We then rely on temporal aggregation of linear time series models to forecast the yearly outcome from the monthly model and hence from monthly information. Relevant papers where this technique is explained are, among others, Amemiya and Wu (1972), Brewer (1973), Weiss (1984), Palm and Nijman (1984), Nijman and Palm (1990), Drost and Nijman (1993). Briefly, this technique allows us to compute the parameters of the annual models, that we use for prediction, from the estimated monthly parameters, that incorporate all the intra-annual available information.

The application of the temporal aggregation technique leads to good predictions for the annual French State deficit. We compare our results with those which would be given by monthly ARIMA models (meaning that, for a given month within the year, we forecast as many periods as remaining months for the end of the year. The annual forecast is the sum of the observed months plus the forecasts) and with the official forecasts made available by the government. It turns out that our predictions are closer to the observed values than monthly ARIMA predictions and the official ones.

The interest of using temporal aggregation techniques emerges very clearly from the improvement in accuracy of the model resulting from the aggregation as compared to the very low accuracy of monthly forecasts. This improvement is due the fact that the aggregation eliminates the errors induced by extraordinary factors, i.e. discretionary measures, which can disrupt time series forecast (the easiest example are years in which the payment of an instalment of the revenue tax is postponed by some months). Moreover, our forecasts are comparable to the official forecasts by the French authorities (they are actually better in some cases and worse in other cases). In fact, the previous literature (for example Bretschneider et al. and Grizzle and Klay quoted above), which compared official predictions based on traditional methods with the outcomes of purely statistical techniques, always found that official predictions were more accurate. The superiority of the official forecast is usually attributed to the fact that official authorities have access to a larger set on information at any time. Future work is needed to compare the accuracy of estimates using temporal aggregation techniques with a combination of statistical and traditional methods, which the existing literature uniformly indicates as the most accurate and trustful.[1]

Our method of temporal aggregation allows to easily update predictions every month, as soon as new intra-annual information is released. A reassuring property of our methodology is that forecast accuracy almost always increases during the year, i.e. with the expansion of the information set upon which the forecast is based on. In addition, while updating a government finance forecast with the traditional method can take time as it requires a complete re-assessment of the macroeconomic situation, the forecasts made using our method can be updated instantaneously. Finally, our method is also fully transparent, as the margin

[1]This resembles the long-lasting debate between the economists that prefer simple statistical-based forecast methods and economists that prefer large models. In general, as far as forecasting properties are concerned, it has been found that simple ARIMA models provide better short run forecast than complex dynamical systems (see, for a very recent example, Marcellino 2004).

of manoeuvre of the forecaster is virtually absent and it can be easily replicated without using much resources.

Of course, the result of the paper and the method we apply are also subject to *a priori* drawbacks. But *a posteriori* they are either solved implicitly by the method or they are out of the scope of the paper. First and most relevant, the temporal aggregation method is subject to criticisms that are inherent to its purely statistical nature. Notably, it does not allow to take into account existing information on discretionary measures that may be taken in the course of the year. We would not be able to incorporate such information in our estimates. However, in this respect, it can be noted that our forecast maintains its validity if the within-year discretionary policy implemented by governments is related to the deficit situation. This seems to us to be the case: in particular, anecdotal evidence points to government intervention when the initial target is likely to be missed by a large margin for instance, and not in normal times. A second, and similar, drawback compared to the traditional approach, is that our methodology provides little economic arguments to explain the results. This is relevant because this forecast have to be communicated to policy makers, who are used to decide on the basis of the economic model which inspires national accounting. However we consider that this criticism does not impair the relevance of our approach because, as far as the accuracy of the forecast is good, the purely statistical forecast can always be used both as a check for official forecast and as an advanced signal on the situation.

A second group of criticisms concerns the type of intra-annual data we can use and has different aspects. First, we are able to apply the method only to the Central government sector (the 'State sector', more precisely). While the relevant variable from a policy point of view is the general government deficit in the institutional setting of EMU, and the State deficit is only a component of the general government deficit. Indeed, we would have preferred to apply our method to a variable covering a broader field, namely the general government sector. This is not possible because of the lack of data, as statistics for developments in general government finances exist only on an annual basis. However, focusing on the State deficit does not diminish the usefulness of our work in the case of France, because the cash State deficit is the largest and most variable component of the general government deficit, and the evolution of the two variables is extremely well correlated. In addition, the nature of the variable we focus on (the monthly State sector deficit, on a cash basis) is different from the variable of interest in the context of EMU (the general government deficit), which is measured using the accrual principle according to usual national accounting principles. Unfortunately, intra-annual national accounting data for the general government are not available. Indeed, although not completely compatible with the national accounting format, the data we focus on are strongly correlated with national accounting figures, as we said before, making them useful prediction tools.

Third, the field covered by the variable we focus on (the State sector cash deficit) is not constant over time, as the precise perimeter of the State sector changes every year and past discretionary measures may make our forecasts less pertinent. This is a relevant issue. We proceed to the necessary econometric analysis to ensure that it was not the case in the past. The outliers detection allows to take into account past changes of the perimeter of the State sector and discretionary measures, both transitory and permanent. In fact, past discretionary measures or changes of perimeters are detected if they imply abnormal departures from the

”normal” behaviour of the time series (e.g. exceptional delays in tax payments). These departures are treated as outliers from an econometric point of view.

The paper is organized as follows. The next two sections discuss data issues, with the aim to clarify not only the quality of the data used, but also to enlighten the difference between annual national accounting and monthly cash data. Section 4 reports the monthly estimations using ARIMA models. Section 5 presents the econometric methodology used in the paper. We do not explain either at length or in technical detail but intuitively, sticking to the main ideas and concepts. We refer the reader to Silvestrini and Veredas (2004) for all the technicalities of the method. Section 6 discusses the results, providing our forecasts for 2002 and 2003, and comparing them with the official French government predictions. Section 7 concludes.

2 Issues related to data

In the European context, the notion of general government deficit is precisely defined. Regulations have been adopted by all members of the European Union in order to build a unified accounting system, the European System of National Accounts (ESA95), with the aim of creating a set of macroeconomic and public finance data which could allow comparisons between the member countries. All legislative acts and economic policy practitioners use figures computed according to these rules. The general government deficit, in the national accounts definition, is therefore the variable of main interest from a policy point of view, and it is the variable which governments would like to be able to estimate.

2.1 Existing data

In principle, it would therefore be suitable to use intra-annual national accounts data to forecast the general government deficit. Unfortunately, at the moment Eurostat - the body of the European Commission in charge of collecting and checking the data estimated according to ESA95 and computed by the various national statistical offices - provides a complete set of non financial data only on an annual basis. Official intra-annual estimates in the ESA95 format at the quarterly frequency exist for the macroeconomic variables, but the complete set of quarterly ESA95 data for the general government deficit and for government finance data in general is still under construction. The lack of intra-annual data within the ESA95 standard implies that a close monitoring of developments in the general government deficit, such as that made for instance to periodically adjust macroeconomic forecasts on the basis of quarterly national accounts released, is not possible.[2]

However, a large amount of monthly information exists on budgetary developments in most countries, in general relative only to certain subsets of the general government. These data are normally part of the sources out of which ESA95 data will be computed at the end of the year, but, as such, they lack many of the interesting properties of the national accounting

[2]One could argue that the improvement in the course of a year of the forecast for the macroeconomic developments will lead to an equivalent improvement in the budgetary forecast for the same year. In practice, the link between macroeconomic developments and budgetary evolutions is far from stable.

data. First of all, only ESA95 data are comparable across European countries; this is of course a very relevant issue in the EU context, where a collective discipline has to be enforced, which requires that any minister is in a position to analyze the government accounts of other countries. Second, there is a difference in coverage, and the infra-annual data have to be (partially) consolidated in order to obtain national accounting figures. Third, these data are calculated on a cash basis, while ESA95 statistics are computed applying the accrual principle.[3] Finally, often, they are not seasonally-adjusted. These elements imply that calendar factors and other specific events may introduce differences between the two sets of figures.

2.2 Reasons for the choice of the monthly cash government deficit as relevant intra-annual data

As stated above, the absence of intra-annual ESA95 government data makes it necessary to choose other intra-annual indicators in order to predict the deficit of the general government in ESA95 format. The choice of which data to use varies across countries depending on the availability and on the quality of existing data and their relevance to forecast the ESA95 general government government deficit. In the case of France, we limit the study to the behaviour of the monthly cash State deficit.

In order to understand precisely the nature of the data we use, it is necessary to explain the relationship between them and the general government deficit. In national accounts, the general government sector is defined as the sum of three big sub-sectors: i) central government, ii) social security and iii) local authorities.

The central government sector is divided in two sub-sectors, i.e. the State sector and the sector of other central government bodies. The State sector consists of all State administration bodies, like the ministries; the other central government bodies include some public firms, museums, state-owned research centres and other similar bodies. The social security sector is a composite of the health sector, unemployment benefits and the pension system. The local authorities sector is composed by all levels of local government. Annual ESA95 data are available not only for the general government sector, but also by sub-sectors. However, as already mentioned, no intra-annual data exist for the general government or its sub-sectors in the ESA95 format.

Other data exist: the length of the series, their availability and their quality depends very much by sub-sector. Only very few intra-annual data are available concerning the sub-sectors of local authorities and of other bodies of central government. In the social security area, cash

[3]The difference between cash and accrual data concerns the time at which the data are recorded. Cash accounting records only cash payments and record them at the time these payments occur. This has the practical advantage that it avoids problems connected with the evaluation of non-monetary flows. Yet, cash accounting cannot be used for economic and national accounting, as the times at which the payments take place may diverge significantly from the economic activity and transaction to which they relate, and it is these underlying activities and transactions that the national accounting system seeks to portray. Accrual accounting records flows at the time economic value is created, transformed, exchanged, transferred or extinguished. This means that flows which imply a change of ownership are entered when ownership passes, services are recorded when provided. The national accounting system favours accrual accounting because the time of accrual accounting is in full agreement with the way economic activities and other flows are defined in the system.

monthly data are available for a part of the health sector. In France, like in most EU countries, the State is the only general government body which provides detailed data on intra-annual budgetary developments. Indeed, every month, the Ministry of Finances publishes a complete set of figures on the evolution of expenditures, tax revenues and the budgetary balance for the State finances. These data are computed on a cash basis.

The data are described more precisely in the next section. For the moment, it is important to understand the scope of these data in comparison with ESA95 central government data. The cash data represent just the expenditures and the revenues perceived by the State every month. In that sense, they are not exactly comparable to the ESA95 data for the State sector, since they contain also revenues and expenditures which are attributed to other sub-sectors, other bodies of central government in particular. The State cash deficit and the State ESA95 deficit for the State sector however follow a parallel pattern.[4] This is even true for most of the revenue and expenditure components considered in the paper. This has led us to limit the study to the State deficit using cash data. We argue that this does not constitute a big limitation, as major forecasting difficulties are concentrated in this sector, and as budgetary developments at the State level largely determine those at the general government level. These intuitions are confirmed by empirical data (see Figure 1). As it is already shown in Perez (2002), a major part of the variation in (ESA95 definition) general government deficit is explained by the variation in the cash State deficit. Forecasting the cash State deficit is thus the most relevant task. It is also the most difficult, given that the State deficit is more variable, as compared to the deficit of other sub-sectors of the general government deficit.

[FIGURE ONE ABOUT HERE]

By contrast, budgetary developments in the social security sector are largely dependent on parameters which are known at the beginning of the year (entitlement rules, for instance), and on economic developments. These evolutions can be forecasted relatively easily. Indeed, while the impact of changes in the macroeconomic environment on general government revenues received by the State is not trivial (notably because of its implications on the tax rates, the difficulties are much bigger for some progressive taxes such as the income tax or the corporate tax), it is clearly lower for revenues of the social security and the local authorities sectors. Finally, due to the strict legal framework,[5] the local authorities sector enjoys a small and relatively stable surplus.

Given that, the predictions of our method, which are based on a pure time series approach, have a better quality. This is not surprising. Often researchers (see for example Marcellino, 2004) have found that simple ARIMA models provide better short run forecast than complex dynamical systems. Most importantly, with our temporal aggregation approach the annual forecasts are updated every month, as soon as new intra-annual information is released.

[4]Of course there are other differences due, for example, to the fact that these are cash data and not accrual data, and that implicit social contributions do not appear in the figures (by definition), but this is not relevant at the moment.

[5]In France, local authorities are basically obliged to respect the golden rule.

3 Data presentation

Each month, the French Ministry of Finance provides a relatively detailed decomposition of the revenue and expenditure sides of French State finances on a cash basis. Our database contains monthly data for the revenue and expenditure cash deficit components (fifteen time series). For each series, we have realizations ranging from June 1995 to December 2003 (103 observations in all).

The revenue side of the State deficit is decomposed in seven items (see Figure 2). VAT represents 34% of total revenue, and it is the most important source of State revenues. The second most important is 'income tax', 17% , followed by 'corporate tax', 11% . The French 'tax on oil products' (TIPP, *Taxe intérieure sur les produits pétroliers*) represents 8% of total State cash revenues. The numerous other direct and indirect taxes collected by the State sector are grouped in a single variable under the headline 'other fiscal revenues', which represents 8% of the total. The 'non-fiscal revenues', which consist notably in the profits made by public bodies and in revenues stemming from specific taxes and social contributions, represent 8% of the total. Finally, the remaining 13% and that we denote by 'others', include transfers of revenues to Local authorities and the European Union and a negligible item called fonds de concours. By convention, the transfers of revenues perceived by the State to Local authorities and the European Union are recorded as a negative revenue of the State.

[FIGURE TWO ABOUT HERE]

On the expenditure side, the most important item is 'wages and pensions' paid by the State, which amounts to 34% of total State expenditure. 'Debt interests payments', 'social interventions' (expenditure made by the state in the social field), and 'military expenditures' represent respectively 14% , 12% and 11% of total State expenditure. Below the 10% there are the 'functioning expenditures' (another sort of current expenditures), with 6% , the 'economic interventions' and the 'other State expenditures', with 9% . Finally, the 'civilian capital expenditures' (public investment financed by the State), with 5%.[6]

[FIGURES THREE AND FOUR ABOUT HERE]

Raw data are clearly nonstationay, as data are released in acumulated terms through the year. Therefore, we base all the subsequent analysis on the their first differences. Figures 3 and 4 show the time series plots for the revenues and expenditures side variables. At first sight, most of the variables present systematic seasonal patterns. This is normal, since data provided by the authorities for the State sector are not seasonally adjusted and largely reflect the calendar of payments. Actually, the two most difficult variables to forecast, corporate and income taxes, display the most systematic seasonal pattern. This is supported by Figure 5, that shows the observations for each month overlapped over the years (on the revenue side). VAT, the most important revenue source, has also a remarkable seasonal pattern, though slightly

[6]There are two additional components, but whose contributions are negligible. They are *Garanties et Degrevements* and *Pouvoirs Publiques*.

more erratic than corporate and income taxes. Other variables, like the tax on oil products, show a more random pattern.

[FIGURES FIVE AND SIX ABOUT HERE]

A similar seasonal analysis is conducted for the expenditure side, and it is presented in Figure 6. Again, a seasonal pattern is evident. As expenditures are largely independent from cyclical fluctuations, which are highly seasonal, this is somewhat counter-intuitive. The seasonality detected here reflects the fact that some expenditures are made according to a precise and constant calendar (interest payments, civil servants bonuses, social interventions corresponding to a payment made by the State to specific to specific populations at a given moment of the year).

[TABLE ONE ABOUT HERE]

Table 1 shows some descriptive statistics. We observe that all the variables have a strong dynamic component, that is they are autocorrelated (see ρ_1 and ρ_2). This is a net indication of the need of time series models to capture correctly the dynamics of these variables. The models should also include a seasonal component, as all the variables are not only correlated with their immediate past, but also with the observations of the same month of previous years (see ρ_{12} and ρ_{24}).

4 Monthly estimations

In order to fully understand the econometrics of this article, we introduce some notation. Let y_t be the monthly random variable, i.e. any of the fifteen revenue or expenditure variables, observed at time t, for $t = 1, \ldots T$. We model the fifteen variables with stationary $ARIMA\,(p,d,q) \times (P,D,Q)_{12}$ processes

$$\Phi\left(L^{12}\right)\phi\left(L\right)\Delta^{d}\Delta^{D}y_t = \Theta\left(L^{12}\right)\theta\left(L\right)\varepsilon_t, \tag{1}$$

where L is the lag operator, i.e. $y_{t-j} = L^j y_t$, $\phi\left(L\right), \theta\left(L\right), \Phi\left(L^{12}\right)$ and $\Theta\left(L^{12}\right)$ are polynomials of order p, q, P and Q respectively, $\Delta^d = (1-L)^d$ and $\Delta^D = (1-L^{12})^D$. The first two polynomials account for the short-run dynamics, while the last two capture the dynamics in seasonality. The left hand side polynomials in (1) are the autoregressive part of the model, while the right hand side are the moving average. D and d are the integration orders. If, for example, $D = d = 0$, y_t is stationary. However, if $d = 0$ and $D = 1$, y_t is stationary in levels but not in seasonality. That is, it has a unit root in seasonality. Finally, the error term ε_t is *i.i.d.* Gaussian with zero mean and constant variance σ^2_ε.

For each revenue and expenditure variable, we estimate an ARIMA model, including an outliers analysis. We use the package TRAMO to fit the best model.[7] Tables 2 and 3 show the

[7]TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) performs estimation, forecasting, and interpolation of regression models with missing observations and ARIMA errors, in the presence of several types of outliers. TRAMO estimates a battery of models, including the outliers analysis, and selects the best model using the Bayes Information Criteria.

results using the whole sample.[8] The first table shows the estimated models, while the second shows the outlier analysis.

[TABLE TWO ABOUT HERE]

All the models have a unit root in seasonality. This is expected, as the seasonal movements of the revenues and expenditures (see Figures 5 and 6) are very remarkable, and they dominate the structure of the time series. Only three of the fifteen models have autoregressive components. And twelve have seasonal moving average components. This means that almost all the variables posses a unit root in seasonality and that the dynamics are dominated by the seasonal errors, that is the errors committed at time $t-12$. Finally, almost all the variables have some outliers. Outliers represent realizations that are associated to special events capable of producing a distortion on the observed time series. The outliers detection is important for two reasons: i) it permits to do correct inference, regardless of the *extreme values* that, otherwise, may introduce some noise into the model, and ii) their detection helps us to understand and to explain specific economic events, like discretionary measures.

As a matter of fact, TRAMO detects three types of outliers. The first, and most common, are additive outliers, that represent an isolated spike. Second, the intervention outliers, which represent a spike that takes a few periods to disappear. The interpretation of these two type of outliers is similar, as their effect is transitory. Finally, TRAMO also detects level shifts, which represent a step function, i.e. a permanent change. Table 3 shows the outliers analysis provided by TRAMO.

[TABLE THREE ABOUT HERE]

Indeed, as we expected, additive and intervention outliers are relatively frequent. The most significant ones correspond to a temporary change of calendar in the collection of a given tax: for instance, the outliers of the income tax for the months of September and October 2002, that can easily been seen in Figure 5, correspond to the fact that the government decided to give an additional delay of one month for the payment of the annual income tax.[9] Most of the other outliers for corporate and income taxes are in April and September respectively. These are the two most important months for each tax as it is when the first instalment is done. And these taxes are characterized by its volatility - see, for example, in Figure 5, the big variations in corporate taxes for in instalment of April. Therefore, the highest fluctuations from the instalments of one year to another are expected in these months. And the unusual fluctuations are captured by the outliers.

Finally, the analysis detects three level shifts in all. The level shift of the tax on oil product (TIPP) of October 2000 corresponds to the introduction of a mechanism aimed at limiting the

[8] This is an important detail. In Section 6 we will show forecasts for different sample sizes. For ease of exposition we only show the estimation results for the complete sample. But the reader should always keep in mind that forecast for different sample sizes may imply different models (for different sample the best model may differ) and hence different estimates.

[9] These outliers will play a major role in Section 6.

impact on consumer prices of large fluctuations in oil prices. This decision was taken in October 2000, when oil prices increased sharply and has led to a significant decline in the yield of this tax (despite the increase in oil prices observed in 2003, the level of this tax in 2003 should still be below its level of 1999), which is detected by the model. The level shift of May 1999 detected on the corporate tax data does not correspond to a specific policy measure. Corporate tax revenues are very volatile, making the identification of outliers particularly difficult. It is recalled that corporate tax revenues increased by 40% between 1997 and 1999 (25% for the sole year 1999). Our model identifies such developments as a change of regime. However, even if discretionary measures were introduced between 1997 and 1999, they do explain only a marginal share of the volatility of the variable. Recent developments tend to confirm this assessment. In the last two years (2002 and 2003), the corporate tax revenues were down by a cumulated 25% and revenues are projected to further decline in 2004. Most probably, the level shift identified by the model over the period 1996-2003 will not prove to be permanent. The reasons for the third structural break identified on the series of other fiscal revenues are even more difficult to explain. This is because this variable is an aggregation of a large number of distinct taxes. However, given the past regularity of the series, it is likely that the new pattern corresponds to the impact of policy decisions.

These are the results of the ARIMA monthly models. Yet our aim is to do annual forecasting, but using all this intra-annual information. To this end, we rely on the techniques of temporal aggregation for ARIMA models.

5 Temporal aggregation: the technique

The aim of this section is to explain intuitively this technique. All the technicalities are skipped and we refer the eager reader to Silvestrini and Veredas (2004) for a rigorous compendium of the method.

The starting points is model (1). First, we estimate the parameters, namely those in the polynomials plus the variance of the error term. Next, we aggregate our variables. Both revenues and expenditures are flow variables, i.e. the annual values are the sum of the 12 months. Let y_t^* be the aggregated variable: $y_t^* = \sum_{j=0}^{11} y_{t-j}$.

We assume that the fifteen components at annual level are represented with $ARIMA\,(p', d', q')$ models

$$\alpha\,(B)\,\Delta^{d'}\,y_T^* = \eta\,(B)\,\varepsilon_T^*, \tag{2}$$

where $T = 1, 2, \cdots$ represents the annual frequency and y_T^* is the sequence of nonoverlapping aggregates of y_t. For instance, $T = 2002$ means the sum of the twelve months of 2002. Equivalently to model (1), $\alpha\,(B)$ and $\eta\,(B)$ are polynomials of orders p' and q', respectively. The lag operator $B = L^{12}$ is now in time unit T, i.e. year, such that $By_T^* = y_{T-1}^*$. The error term is *i.i.d.* Gaussian with zero mean and variance $\sigma^2_{\varepsilon^*}$. Notice the similitude and differences between models (1) and (2). They are both ARIMA models, but model (2) does not have seasonal polynomials or seasonal unit roots. Since the aggregation frequency is 12, the intra-annual seasonality vanishes at annual level.

The core part of the technique is to compute the correct orders p', d' and q', as well as the parameters in $\alpha(B)$, $\eta(B)$ and the variance $\sigma^2_{\varepsilon^*}$, from the parameters and polynomial orders of the monthly model that contain all the monthly information.

The first step is to obtain the orders p', d' and q'. They are computed using the following relations

$$\begin{aligned} p^{'} &= p + P \\ d^{'} &= d + D \\ q^{'} &= \left\lfloor \frac{1}{12} \left(11\left(p+d+1\right)+q+12Q\right) \right\rfloor, \end{aligned} \tag{3}$$

where $\lfloor\cdot\rfloor$ is the integer part of the argument.[10]

Second, the autoregressive parameters, $\alpha(B)$, are computed from their monthly counterparts, $\phi(L)$ and $\Phi\left(L^{12}\right)$. Denote $\delta_i, i = 1, \ldots, p+P$ the inverted roots of the polynomial $\phi(L)\Phi\left(L^{12}\right)$, i.e. $\phi(L)\Phi\left(L^{12}\right) = \prod_{i=1}^{p+P}\left(1-\delta_i L\right)$. Let

$$T(L) = \prod_{i=1}^{p+P}\left[\frac{1-\delta_i^{12}L^{12}}{1-\delta_i L}\right]\left[\frac{1-L^{12}}{1-L}\right]^d \tag{4}$$

be a polynomial that links the monthly and the annual autoregressive parameters. We establish the following equality

$$\alpha(B) = \phi(L)\Phi\left(L^{12}\right)T(L). \tag{5}$$

Everything in the right hand side is known. The polynomial order, p, the integration order, d, the parameters in $\phi(L)$ and $\Phi\left(L^{12}\right)$ and their inverted roots $\delta_i, i = 1, \ldots, p+P$, are already specified and estimated.

For example, suppose that the monthly model in (1) is an AR(1): $y_t = \phi y_{t-1} + \varepsilon_t$. Then

$$\begin{aligned} T(L) &= \frac{1-\phi^{12}L^{12}}{1-\phi L}, \\ \alpha(B) &= (1-\phi L)\frac{1-\phi^{12}L^{12}}{1-\phi L} = (1-\phi^{12}B). \end{aligned}$$

That is, the annual model is an $AR(1)$ with parameter $\alpha = \phi^{12}$. If the monthly model is stationary, α is very small. When the aggregation frequency is high, the AR parameters become unimportant. This makes sense as they capture the short run memory of the process. When we aggregate this memory is lost.

[10] These equations, as all in this section, are not general but specific to the case when the seasonal frequency equals the aggregation frequency. In our case both are equal to 12. A counter-example is if we aggregate monthly into quarterly. The aggregation frequency is 3, but the seasonal frequency is 4. The general expressions are quite unhandy and worthless for the scope of the paper. But they may be found in Silvestrini and Veredas (2004).

Finally, the moving average parameters, $\eta(B)$, are computed using the variance-covariance structures of $\theta(L)$ and $\eta(B)$. Unfortunately, there are no general expressions to get these parameters. In the remainder of this section we show how to obtain them with two examples that cover a range of the models encountered in our empirical application.

5.1 Example one: $ARIMA\,(0,0,0) \times (0,1,1)_{12}$

The monthly model, with constant, is

$$\left(1 - L^{12}\right) y_t = c + \left(1 + \Theta L^{12}\right) \varepsilon_t. \tag{6}$$

First, we express this model in annual frequency. We multiply both sides of (6) by $\sum_{i=0}^{11} L^j$

$$\left(1 - L^{12}\right) y_t^* = \sum_{j=0}^{11} L^j c + \left(1 + \Theta L^{12}\right) \sum_{j=0}^{11} \varepsilon_{t-j}. \tag{7}$$

This model is used to compute the order $d^{'}$, the parameters in $\alpha\,(B)$, $\eta(B)$ and the variance $\sigma^2_{\varepsilon^*}$. Using (3), the AR and MA polynomials are of order $p^{'} = 0$ and $q^{'} = \lfloor (1/12)\, 23 \rfloor = 1$. The integration order is $d^{'} = 1$. Therefore, the annual model is an $ARIMA(0,1,1)$ with constant:

$$(1 - B)\, y_T^* = C + (1 + \eta B)\, \varepsilon_T^*. \tag{8}$$

We now compute the aggregated parameters. Comparing (7) and (8), the constant term, C, is straightforward, $C = 12c$. For the moving average parameters and $\sigma^2_{\varepsilon^*}$, we equate the variance-covariance structures of models (6) and (8). As they are both moving averages of order one, only the variance and the autocovariance of order one are different from zero. From the monthly model (6), the variance is $\gamma_0 = 12\left(1 + \Theta^2\right)\sigma^2_{\varepsilon}$ and the autocovariance of order one is $\gamma_1 = 12\Theta\sigma^2_{\varepsilon}$. For the annual model (8), the variance is $\Gamma_0 = \left(1 + \eta^2\right)\sigma^2_{\varepsilon^*}$ and the autocovariance of order one is $\Gamma_1 = \eta\sigma^2_{\varepsilon^*}$. The parameters in γ_0 and γ_1 are already estimated and hence γ_0 and γ_1 are constants. Equating $\gamma_0 = \Gamma_0$ and $\gamma_1 = \Gamma_1$, we get a system of two equations with two unknowns which is easily solved.

5.2 Example two: $ARIMA\,(0,0,1) \times (0,1,1)_{12}$

The monthly model, with constant, is:

$$\left(1 - L^{12}\right) y_t = c + (1 + \theta L)\left(1 + \Theta L^{12}\right) \varepsilon_t. \tag{9}$$

As before, we express it in terms of the aggregate frequency. We multiply both sides by $\sum_{i=0}^{11} L^j$, obtaining

$$\left(1 - L^{12}\right) y_t^* = \sum_{j=0}^{11} L^j c + (1 + \theta L)\left(1 + \Theta L^{12}\right) \sum_{j=0}^{11} \varepsilon_{t-j}, \tag{10}$$

and this aggregated model corresponds to an $ARIMA(p^{'}, d^{'}, q^{'})$, where $p^{'} = 0$, $d^{'} = 1$ and $q^{'} = \lfloor (1/12)\, 24 \rfloor = 2$,

$$(1 - B)\, y_T^* = C + \left(1 + \eta_1 B + \eta_2 B^2\right) \varepsilon_T^*, \tag{11}$$

where $C = 12c$. We compute the parameters η_1, η_2 and $\sigma^2_{\varepsilon^*}$ equating the variance-covariance matrices of ε_t and ε^*_{12t}. For the monthly model:

$$\begin{aligned}
\gamma_0 &= [12 + 12\theta^2 + 12\Theta^2 + 12\theta^2\Theta^2 + 2\theta(11 + \Theta + 11\Theta^2)]\sigma^2_\varepsilon \\
\gamma_1 &= \left[(\theta + \Theta) + 11\,(1+\theta)^2\,\Theta + \theta\Theta\,(\theta + \Theta)\right] \sigma^2_\varepsilon \\
\gamma_2 &= \theta\Theta\sigma^2_\varepsilon.
\end{aligned} \tag{12}$$

Note, again, that all the parameters are already estimated and hence these expressions are constants. For the annual model:

$$\begin{aligned}
\Gamma_0 &= \left(1 + \eta_1^2 + \eta_2^2\right) \sigma^2_{\varepsilon^*} \\
\Gamma_1 &= (\eta_1 + \eta_1\eta_2)\, \sigma^2_{\varepsilon^*} \\
\Gamma_2 &= \eta_2\sigma^2_{\varepsilon^*}.
\end{aligned} \tag{13}$$

Equating (12) and (13) we get a system of three equations in three unknowns.

6 Temporal aggregation: results

6.1 Annual models

We apply the methodology explained in the previous section to our fifteen revenue and expenditure components. The aggregated models are reported in Table 4. All the annual models have, at least, one unit root. In effect, the seasonal unit roots in the monthly models become ordinary unit roots at annual level, as seasonality has vanished. The AR order is unchanged by temporal aggregation. The parameters of the AR polynomial of the annual series model are the 12th powers of the monthly AR polynomial and, consequently, they become very small in value. All the models include some moving average of orders one or two. In sum, the AR components become unimportant, the number of unit roots remains unchanged and the moving average components simplify but remain relevant. Everything is consistent with the theory.[11]

[TABLE FOUR ABOUT HERE]

[11] A note about standard deviations: although not reported, standard deviations of the annual parameters are theoretically possible to compute. This is done using the Delta method. For AR parameters it is straightforward. However, for the MA parameters it is slightly more difficult, as the annual MA parameters are the result of the nonlinear system of equations.

The aggregated models do not incorporate the outliers. This is not a crucial issue for three reasons: i) at annual frequency, the intra-annual outliers lose virtually their effect, ii) although we may apply the aggregation technique to the outliers, their parameters tend quickly to very small values and iii) for annual forecasting the aggregation of outliers does not add any valuable information.

6.2 Forecasting

We assess the predictive performance of our method for the last two years, 2002 and 2003. Although models shown previously have used the whole sample, for the forecasting exercise we use the appropriate in-sample. For instance, starting in January 2002, we estimate a monthly model for each component and we make a prediction for the whole year. Next, we move one month ahead and we re-estimate the models including the February observation, and then we update the annual forecast. This process continues up to December, transferring progressively observations from the out-sample to the in-sample. Two different forecasting methods are applied.

1. *Monthly cumulation*: We only make use of the monthly models. For a given month, we forecast as many ahead periods as the remaining number of months to the end of the year. For example, using observations up to April 2002, we fit a monthly model and we predict eight periods. The annual forecast is the sum of the four observations up to April plus the eight predictions up to December.

2. *Annual aggregation*: We use the temporal aggregation technique. Every month, the estimated models are temporally aggregated into the corresponding annual models. These are used for the one step ahead annual forecasting. For instance, using information up to April 2002, we fit a monthly model, we aggregate it and we forecast at annual level for 2002. The information within 2002 is summarized in the parameters, that are re-estimated every month.

Both procedures share an important feature: as soon as new intra-annual information is released, the estimated models are promptly updated. Exactly the same exercise is repeated for 2003. Tables 5, 6 and 7 report the forecasts for all the revenues and expenditure components for 2002 and 2003. For the sake of exposition we only show four predictions, based on four different information sets, namely, information up to January, April, July and October. These are the interesting months from an institutional point of view: January has the payment of the most important VAT instalment. April is the month in which the French authorities publish their first intra-annual forecasts for the general government deficit. In addition, the amount collected with the payment of the first instalment of the corporate tax is known; April is also a relevant month for VAT in France, because reimbursements of the VAT of the previous year are made up to March, so that April is the first "clean" data on VAT. On the expenditure side, the Government has already spent an important fraction of the total annual expenditures (like debt interest payments or wages and pensions) at this time of the year. July is the month in which the macroeconomic forecast for the next year budget is started and October is the month in which the next year draft budget bill is presented to Parliament. In this context, every October, the French Ministry of Finance releases official predictions on the State (and general government) deficit for the current and next year. For instance, in October 2002, the French government released predictions for the final yearly outcome for 2002 (an update of the forecasts made in October 2001) and 2003. We report the official forecasts by the French government. The forecasts for the revenue variables made by the French authorities should incorporate the largest possible number of information, because during the year the French authorities have at any point access (at least partial) to information on the actual behaviour of revenues and expenditures and to the reasons of divergence of it from the forecasted behaviour.

[TABLE FIVE, SIX, SEVEN ABOUT HERE]

Finally, Figures 7 and 8 report the forecasted State deficit for 2002 and 2003 using our method. These predictions are computed as the difference between the sum of all the revenues forecasts and the sum of the expenditures forecasts. They include hence the forecasts of Tables 5, 6 and 7. These Figures also show the official predictions.[12] Three main conclusions can be drawn from Tables 5, 6 and 7 and Figures 7 and 8.

First, in general our predictions, in particular those based on temporal aggregation, converge toward the observed values as new intra-annual information is used. For instance, a detailed view to VAT (which represents 40% of the total state revenues) reveals that the official forecast in October 2001 for 2002 is above the observed value, 111.3 against 107.5. Our predictions with information up to January are also upward biased, 108.5, though less than the official ones. However, in April 2002, updating our forecast with the new information, we are able to give already a very accurate prediction (107.82). Similar conclusions may be drawn regarding almost all the other components. This is very relevant in view of the possibility of using this technique as an early signal in the course of the year about the situation of the state budget: the results from updating the exercise in the course of the year correctly indicate whether the actual deficit will tend to be in line with the official forecast or not. This is interesting in particular to test the fact that it is worth to pass from the simple ARIMA predictions to the aggregation technique that we propose.

[FIGURES SEVEN AND EIGHT ABOUT HERE]

Second, a glance at Figure 7 shows that the technique based on temporal aggregation outperforms the official forecasts of the State deficit, whatever information set we consider. For example, in October 2001 the French government released a forecast of the State deficit for 2003 of -32.05 billion. The next official update publicly available was given one year later, in October 2002, three months before the end of the year. However, with our method in April and July 2002 our prediction showed already a serious departure from the official one. And, indeed, it was very close to the final realization. The large change with respect to January prediction is due to the important amount of information content in March and April.

[12]Instead summing all the forecasts, we could use techniques of cross-section aggregation, that is when we aggregate models of different variables in the same moment of time. However, this is, first, out of the scope of this paper and, second, it needs some theoretical research as it is not as well developed as temporal aggregation.

As previously said, April is a very important month as payments of the first instalment for some taxes are know and an important share of the annual expenditure is done within the first quarter. Subsequent predictions in incoming months (July and October) supported this difference with the official numbers. Consequently, by the beginning of the second quarter, we were able to forecast the worsening of the State deficit, that finally happened at the end of the year. Moreover, as from April 2002, and even more clearly in July of that year, the forecast based on the temporal aggregation technique already foresaw the large slippage in expenditure that occurred in 2002, as well as the significant shortfall in tax revenues: the results from the aggregation technique are at least comparable to the official forecast by the French authorities, a result which constitutes an improvement with respect to the existing literature. The results for 2002 are very encouraging, because in this year the actual forecast of the French deficit were hotly debated in the European and national context, given that the EU started the excessive deficit procedure in the case of France and, in the same year, there were political elections in France.

The same goes for 2003, but with some reservation. If we compare our forecasts in April and July with the official ones in October 2002, we reach the same advantageous conclusions. That is, as the year goes on we can forecast the worsening of the deficit. However, comparing with the forecast of October 2003, the analysis by component shows that the official results released in October 2002 were at least comparable with the ones from the application of the aggregation technique, so that the accuracy in the deficit forecast comes from the contemporaneous presence of two errors that compensate each other. However, the other property of the model still holds and the official predictions do are really more accurate only in some item. And results are only comparable from the revenue side, as there are no official forecasts for the expenditures. In sum, this forecasting exercise has been proved to be capable of reasonably well identifying deficit deviations early in the year. Our forecasts are moreover comparable to the official ones. This constitutes an improvement compared to the previous studies quoted in the introduction, which generally found a clear superiority for "traditional" forecasting methods.

Finally, the annual aggregation outperforms the monthly cumulation. This is particularly evident for the State deficit forecast for 2002 (see Figure 7). The cumulation method, indeed, forecasts in October 2002 a final deficit of 64.95, against a final realized value of 48.02. The forecast we obtain with the aggregation method, on the contrary, is much closer to the observed value, 48.64. This difference is due to the effect of the discretionary measures in the monthly predictions. The income tax prediction is largely underestimated by the monthly cumulation method (for October 2002, a forecast value of 41.38 against a realization of 49.99). Although not reported, for any information set, the monthly forecasts for October are always very close to 13 (12.9, 13.1 and 13.0 with information up to January 2002, April 2002 and July 2002, respectively), while the realization is much lower: 1.14. This is due to a discretionary measure already mentioned in Section 4. In September and October 2002 the government exceptionally decided to give an additional delay of one month for the payment of the income tax. This delay can be also clearly seen in Figure 5. Using the temporal aggregation, the influence of this unexpected delay is mitigated. As the periodicity of the aggregated models is always one year, a transitory discretionary measure (like the one taken in September-October 2002), that only changes the calendar at which taxes are collected, has a minor impact on the annual prediction. This has the important implication that updated forecasts using the temporal aggregation technique are robust, although not independent, to unexpected and temporary discretionary measures. This is, in our view, one of the most relevant contribution to the literature, because we show (even if this conclusion is based only on relatively short time series) that the technique of temporal aggregation improves the results compared with standard ARIMA modelling.

7 Conclusion

This analysis is motivated by the need of monitoring fiscal variables in the Euro area, mainly due to the Stability and Growth Pact deficit limits. More precisely, we tried to build an early warning signal indicator for assessing the development of public finances in the short run, exploiting all available intra-annual information deriving from monthly cash deficit components.

The application is conducted focusing on the French cash State deficit (*solde général d'execution*), looking at its components, from the revenue and expenditure sides. For each component, we estimate monthly ARIMA models that are temporally aggregated to the annual frequency, as we are interested in yearly predictions. More precisely, starting from monthly observations from the fifteen deficit components, we first estimate ARIMA models for each component. Once the corresponding parameters are estimated, the second step consists in temporally aggregating each of the monthly models previously estimated, deriving a new parameterization for observations generated at a different periodicity (yearly). The third step is to infer the annual model (i.e. the order of the ARIMA model plus the parameters), and perform yearly predictions for each component consolidating, finally, all these predictions in a single one, for the *solde général d'execution* deficit.

The short-run forecasting exercises performed for years 2002 and 2003 highlight that the one-step ahead predictions based on the aggregated models outperform the official forecasts made available by the French government, for each of the fifteen components and thus for the whole State deficit. More important, by the middle of the year we provide very accurate predictions for the current year. The method we propose could be extremely useful to provide policy makers with an indicator to employ when assessing the development of public finances in the short run (one year horizon or even less). All this stresses the relevance of the approach followed, that appears to be reliable on a theoretical ground and also on a more applied framework.

References

[1] Amemiya, T., Wu, R.Y. (1972). "The effect of aggregation on prediction in the autoregressive model", *Journal of the American Statistical Association*, 67, 628-632.

[2] Box, G.E.P., Jenkins, G.M., Reinsel, G.C. (1994). Time series analysis: forecasting and control, 3rd edition, Prentice Hall.

[3] Bretschneider, S., W. Gorr, G. Grizzle, and W.E. Klay (1989). "Political and organizational influences on the accuracy of forecasting state government revenues", *International Journal of Forecasting*, 5, 307-319

[4] Brewer, K.R.W. (1973). "Some consequences of temporal aggregation and systematic sampling for Arma and Armax models", *Journal of Econometrics*, 1, 133-154.

[5] Camba-Mendez, G. and A. Lamo (2002). "Short-term monitoring of fiscal policy discipline", European Central Bank, Working Papers n. 152.

[6] Drost, F.C. and Nijman, T.E. (1993). "Temporal aggregation of Garch processes", *Econometrica*, 61, 909-927.

[7] Fullerton, T.M. (1989). "A composite approach to forecasting state governments revenues: case study of the Idaho sales tax", *International Journal of Forecasting*, 5, 373-380.

[8] Grizzle, G.A., Klay, W.M. (Fall 1994): "Forecasting state sales tax revenues: Comparing the accuracy of different methods", *State and Local Government Review*, Vol. 26, No. 3.

[9] Kinnunen, H. (1999). "Cash data for central government: time series analysis", Bank of Finland, Economics Department Working Papers.

[10] Lawrence, K., Anandarajan, A., Kleinman, G. (1998). "Forecasting state tax revenues: a new approach", *Advances in Business and Management Forecasting*, 2, 157-170.

[11] Marcellino, M. (2004). "Forecast Pooling for European Macroeconomic Variables", *Oxford Bulletin of Economic and Statistics*, 66, 91-112.

[12] Nazmi N., and J.H. Leuthold (1988). "Forecasting economic time series that require a Power Transformation: the case of tax receipts", *Journal of Forecasting*, 7, 173-184.

[13] Nijman, T.E. and Palm, F.C. (1990). "Parameter identification in Arma processes in the presence of regular but incomplete sampling", *Journal of Time Series Analysis*, 11, n. 3, 239-248.

[14] Pegregal D. and J.J. Perez (2003). "A synthetic leading fiscal indicator for the euro area". CentrA, Mimeo.

[15] Pegregal D., J.J. Perez, and J. Toro (2002). "Leading fiscal indicator for the euro area government deficit". CentrA, Mimeo.

[16] Perez, J.J. (2002): "Forecasting general government deficit in the euro area: the role of fiscal leading indicators", CentrA working paper, Spain.

[17] Palm, F.C. and Nijman, T.E. (1984). "Missing observations in the dynamic regression model", *Econometrica*, 52, n. 6, 1415-1435.

[18] Silvestrini, A. and Veredas, D. (2004). "Temporal Aggregation of Linear Time Series Model: A synthesis and an Application". Mimeo.

[19] Weiss, A. (1984). "Systematic sampling and temporal aggregation in time series models", *Journal of Econometrics*, 26, 271-281.

Variables	**Mean**	**S.d.**	ρ_1 and (Q_1)	ρ_2 and (Q_2)	ρ_{12} and (Q_{12})	ρ_{24} and (Q_{24})
REVENUES						
Income tax	4.12	3.21	-0.25	-0.34	0.66	0.64
			(6.35)	(18.49)	(102.4)	(204.2)
Corporate tax	2.62	3.54	-0.41	-0.28	0.89	0.76
			(18.01)	(26.22)	(302.7)	(568.9)
Tax on oil products	1.97	0.17	-0.090	0.30	0.58	0.18
			(0.85)	(10.62)	(58.74)	(84.09)
VAT	8.39	1.34	-0.16	-0.21	0.81	0.65
			(2.75)	(7.37)	(271.7)	(503.2)
Other fiscal revenues	1.92	1.15	0.23	-0.09	0.67	0.60
			(5.40)	(6.23)	(116.6)	(216.8)
Non fiscal revenues	2.22	1.64	-0.18	-0.12	0.32	-0.021
			(3.30)	(4.82)	(24.64)	(34.44)
Others	-2.91	1.44	-0.22	0.21	0.25	0.12
			(4.83)	(9.35)	(40.91)	(55.40)
EXPENDITURES						
Debt interest payments	3.08	3.84	-0.24	-0.35	0.87	0.73
			(6.12)	(19.01)	(245.8)	(451.4)
Wages and pensions	7.30	0.65	0.51	0.56	0.64	0.36
			(26.82)	(60.17)	(380.9)	(517.6)
Functioning expenditures	1.31	0.51	-0.30	-0.20	0.55	0.46
			(9.46)	(13.55)	(80.00)	(132.4)
Economic interventions	1.98	0.61	0.17	0.034	0.31	0.29
			(3.07)	(3.19)	(44.97)	(76.19)
Social interventions	2.47	0.71	-0.059	-0.21	0.53	0.44
			(0.37)	(4.93)	(47.79)	(87.53)
Other interventions	1.78	0.60	0.29	0.21	0.58	0.45
			(9.05)	(13.65)	(75.77)	(134.8)
Civilian capital exp.	1.33	0.62	-0.21	-0.14	0.63	0.61
			(4.61)	(6.60)	(81.57)	(148.1)
Military expenditures	2.32	0.56	0.006	-0.10	0.49	0.28
			(0.004)	(1.12)	(36.02)	(52.78)

Top part of the table is for revenue side variables, bottom part for expenditure side variables. ρ_i and (Q_i) represent, respectively, the autocorrelation of order i and the Q-stat. of order i. 'Civilian capital exp.' stands for civilian capital expenditures.

Table 1: Descriptive statistics

Variables	**Estimated Monthly Models**	$\hat{\sigma}_\varepsilon^2$
REVENUES		
Income tax	$\Delta\Delta^{12}y_t = -0.00026 + (1-0.94L)(1-0.50L^{12})\varepsilon_t + out_t$	0.1563
Corporate tax	$(1-0.077L)\,\Delta^{12}y_t = \varepsilon_t + out_t$	0.1515
Tax on oil products	$\left(1+0.41L+0.29L^2\right)\Delta^{12}y_t = \varepsilon_t + out_t$	0.0069
VAT	$\Delta^{12}y_t = 0.21 + (1-0.98L^{12})\varepsilon_t$	0.1488
Other fiscal revenues	$\Delta^{12}y_t = -0.019 + (1+0.15L)(1-0.46L^{12})\varepsilon_t + out_t$	0.2053
Non fiscal revenues	$\Delta\Delta^{12}y_t = (1-0.97L)(1+0.11L^{12})\varepsilon_t + out_t$	0.7773
Others	$\Delta^{12}y_t = -0.26 + (1-0.85L^{12})\varepsilon_t + out_t$	0.1868
EXPENDITURES		
Debt interest payments	$\Delta\Delta^{12}y_t = (1-0.76L)(1-0.23L^{12})\varepsilon_t + out_t$	0.6084
Wages and pensions	$(1-0.27L)\,\Delta^{12}y_t = 0.21 + (1+0.24L)(1-0.64L^{12})\varepsilon_t + out_t$	0.0436
Functioning expenditures	$\Delta^{12}y_t = 0.042 + (1-0.61L)(1-0.71L^{12})\varepsilon_t + out_t$	0.0766
Economic interventions	$\Delta\Delta^{12}y_t = -0.0056 + (1-0.92L)(1-0.63L^{12})\varepsilon_t + out_t$	0.1584
Social interventions	$\Delta^{12}y_t = 0.052 + (1-0.51L)(1-0.61L^{12})\varepsilon_t + out_t$	0.1653
Other interventions	$\Delta^{12}y_t = 0.11 + (1-0.50L^{12})\varepsilon_t + out_t$	0.0974
Civilian capital expenditures	$\Delta^{12}y_t = (1-0.30L)(1-0.58L^{12})\varepsilon_t + out_t$	0.0841
Military expenditures	$\Delta^{12}y_t = 0.030 + (1+0.19L)(1-0.27L^{12})\varepsilon_t$	0.2451

Top part of the table is for revenue side variables, bottom part for expenditure side variables.

Table 2: Estimated monthly models

Variables	**Additive Outliers**	**Intervention Outliers**	**Level Shifts**
REVENUES			
Income tax	09/1997 03/2001 09/2002 10/2002	09/1999 09/2000	Not detected
Corporate tax	12/1999 04/2001 06/2001	12/1997 04/1999 04/2003	05/1999
Tax on oil products	Not detected	12/1998 01/1999	10/2000
VAT	Not detected	Not detected	Not detected
Other fiscal revenues	Not detected	10/2000 10/2002	08/2003
Non fiscal revenues	11/2000	12/2000 12/2001	Not detected
Others	05/1997 12/1999 12/2000	04/1997 11/2000 11/2001 11/2002	Not detected
EXPENDITURES			
Debt interest payments	03/1998	Not detected	Not detected
Wages and pensions	Not detected	06/2002	Not detected
Functioning expenditures	07/1995	Not detected	Not detected
Economic interventions	12/1995 03/2001 07/2002	04/2001	Not detected
Social interventions	12/2001 09/2002	Not detected	Not detected
Other interventions	Not detected	02/2001 03/2001	Not detected
Civilian capital expenditures	04/1997	Not detected	Not detected
Military expenditures	Not detected	Not detected	Not detected

Top part of the table is for revenue side variables, bottom part for expenditure side variables.

Table 3: Outliers analysis

Variables	Aggregated Models	$\hat{\sigma}^2_{\varepsilon^*}$
REVENUES		
Income tax	$\Delta^2 y^*_T = -0.037 + (1 - 0.97B + 0.23B^2)\varepsilon^*_T$	3.4277
Corporate tax	$\left(1 - 4.34 \cdot 10^{-14}B\right) \Delta y^*_T = (1 + 0.0065B)\varepsilon^*_T$	2.1063
Tax on oil products	$\left(1 - 2.26 \cdot 10^{-5}B - 3.54 \cdot 10^{-7}B^2\right) \Delta y^*_T = (1 - 0.097B - 2.6 \cdot 10^{-6}B^2)\varepsilon^*_T$	0.0308
VAT	$\Delta y^*_T = 2.52 + (1 - 0.98B)\varepsilon^*_T$	1.7856
Other fiscal revenues	$\Delta y^*_T = -0.23 + (1 - 0.45B - 0.0044B^2)\varepsilon^*_T$	3.1962
Non fiscal revenues	$\Delta^2 y^*_T = (1 - 0.58B - 0.076B^2)\varepsilon^*_T$	12.777
Others	$\Delta y^*_T = -3.12 + (1 - 0.85B)\varepsilon^*_T$	2.2416
EXPENDITURES		
Debt interest payments	$\Delta^2 y^*_T = (1 - 0.14B - 0.02B^2)\varepsilon^*_T$	2.7094
Wages and pensions	$\left(1 - 1.50 \cdot 10^{-7}B\right) \Delta y^*_T = 3.45 + (1 - 0.61B - 0.022B^2)\varepsilon^*_T$	1.4121
Functioning expenditures	$\Delta y^*_T = 0.50 + (1 - 0.92B + 0.15B^2)\varepsilon^*_T$	0.2235
Economic interventions	$\Delta^2 y^*_T = -0.81 + (1 - 0.98B + 0.22B^2)\varepsilon^*_T$	4.1575
Social interventions	$\Delta y^*_T = 0.62 + (1 - 0.74B + 0.081B^2)\varepsilon^*_T$	0.6337
Other interventions	$\Delta y^*_T = 1.32 + (1 - 0.50B)\varepsilon^*_T$	1.1688
Civilian capital exp.	$\Delta y^*_T = (1 - 0.62B + 0.027B^2)\varepsilon^*_T$	0.4895
Military expenditures	$\Delta y^*_T = 0.36 + (1 - 0.26B - 0.0031B^2)\varepsilon^*_T$	4.0714

Top part of the table is for revenue side variables, bottom part for expenditure side variables. 'Civilian capital exp.' stands for civilian capital expenditures.

Table 4: Annual models temporally aggregated (from monthly frequency)

Table 5: Forecasting table for revenue variables.

		Income tax	Corp. tax	Tax oil prod.	VAT	Other fiscal rev.	Non fiscal rev.	Others
	Observed 2002	49.99	37.52	23.96	107.50	21.25	32.76	-44.90
Sample horizon	Forecast							
October 2001	Official	54.05	40.08	24.09	111.30	20.88	34.80	-30.06
January 2002	Cumulated	50.91	42.68	24.25	109.47	21.87	22.59	-44.86
January 2002	Aggregated	52.44	37.01	25.82	108.04	21.99	32.27	-37.53
April 2002	Cumulated	53.12	40.25	24.01	107.87	22.51	22.42	-44.70
April 2002	Aggregated	52.78	37.21	25.17	107.92	22.06	32.25	-41.82
July 2002	Cumulated	52.24	38.51	23.90	107.90	22.76	23.11	-44.59
July 2002	Aggregated	52.73	37.63	24.71	107.92	22.09	32.33	-45.16
October 2002	Official	51.42	37.42	24.35	107.60	21.66	32.35	-45.32
October 2002	Cumulated	41.38	37.75	23.93	106.58	23.33	22.71	-44.15
October 2002	Aggregated	52.78	37.61	24.30	107.80	22.09	32.25	-45.54
	Observed 2003	53.75	35.05	24.30	109.01	17.71	30.16	-48.42
Sample horizon	Forecast							
October 2002	Official	53.00	37.36	25.80	112.02	20.96	31.07	-43.50
January 2003	Cumulated	47.80	37.52	24.45	111.41	21.92	33.78	-49.84
January 2003	Aggregated	53.01	37.52	23.96	110.26	21.16	33.74	-41.65
April 2003	Cumulated	54.87	36.47	24.70	109.67	20.14	32.22	-48.43
April 2003	Aggregated	53.19	39.91	24.51	110.14	21.15	32.76	-45.73
July 2003	Cumulated	54.70	32.78	24.82	108.88	19.56	30.60	-48.12
July 2003	Aggregated	53.16	38.45	24.99	110.02	21.02	31.21	-45.51
October 2003	Official	53.05	34.56	24.67	109.80	19.20	29.70	-53.39
October 2003	Cumulated	54.30	35.35	24.36	108.60	11.69	30.98	-48.24
October 2003	Aggregated	53.42	37.52	23.96	110.02	21.02	31.21	-48.44

Table 6: Forecasting table for expenditure variables (1)

		Social interv.	**Other interv.**	**Civilian capital expend.**	**Military expend.**
	OBSERVED 2002	32.92	26.1	16.81	29.21
SAMPLE HORIZON	FORECAST				
January 2002	Cumulated	32.26	22.04	15.37	27.69
January 2002	Aggregated	31.90	31.20	16.30	27.85
April 2002	Cumulated	31.67	24.61	16.32	28.92
April 2002	Aggregated	31.00	37.03	16.30	27.99
July 2002	Cumulated	30.74	26.08	17.50	29.66
July 2002	Aggregated	30.83	34.42	15.73	28.09
October 2002	Cumulated	31.22	25.82	17.05	30.07
October 2002	Aggregated	30.88	34.37	16.60	28.15
	OBSERVED 2003	32.35	26.94	17.00	30.17
SAMPLE HORIZON	FORECAST				
January 2003	Cumulated	32.36	26.95	17.07	29.05
January 2003	Aggregated	33.62	27.41	17.11	29.37
April 2003	Cumulated	31.31	25.21	16.55	30.52
April 2003	Aggregated	33.52	33.77	17.06	29.55
July 2003	Cumulated	31.49	30.87	17.24	30.57
July 2003	Aggregated	33.53	33.56	17.12	29.55
October 2003	Cumulated	31.64	27.42	17.61	30.71
October 2003	Aggregated	33.54	27.42	16.81	29.57

Table 7: Forecasting table for expenditure variables (2)

		Debt interest payments	Wages and pensions	Functioning expend.	Economic interv.
	Observed 2002	38.05	95.79	17.50	19.71
Sample horizon	Forecast				
January 2002	Cumulated	37.46	94.32	16.80	16.12
January 2002	Aggregated	36.68	94.34	16.84	20.97
April 2002	Cumulated	36.51	94.63	16.94	21.53
April 2002	Aggregated	37.86	94.38	16.86	21.74
July 2002	Cumulated	37.38	99.18	17.28	19.82
July 2002	Aggregated	36.68	94.39	16.89	21.93
October 2002	Cumulated	38.24	95.50	16.96	20.10
October 2002	Aggregated	36.68	94.41	16.85	21.66
	Observed 2003	37.57	97.32	17.54	17.88
Sample horizon	Forecast				
January 2003	Cumulated	38.43	97.80	17.41	15.99
January 2003	Aggregated	39.08	98.42	17.97	19.90
April 2003	Cumulated	38.48	97.69	17.79	17.20
April 2003	Aggregated	41.97	98.36	18.00	20.58
July 2003	Cumulated	38.18	97.85	17.71	17.10
July 2003	Aggregated	38.42	98.36	17.99	20.50
October 2003	Cumulated	37.81	97.83	17.80	17.07
October 2003	Aggregated	36.68	103.99	18.00	20.32

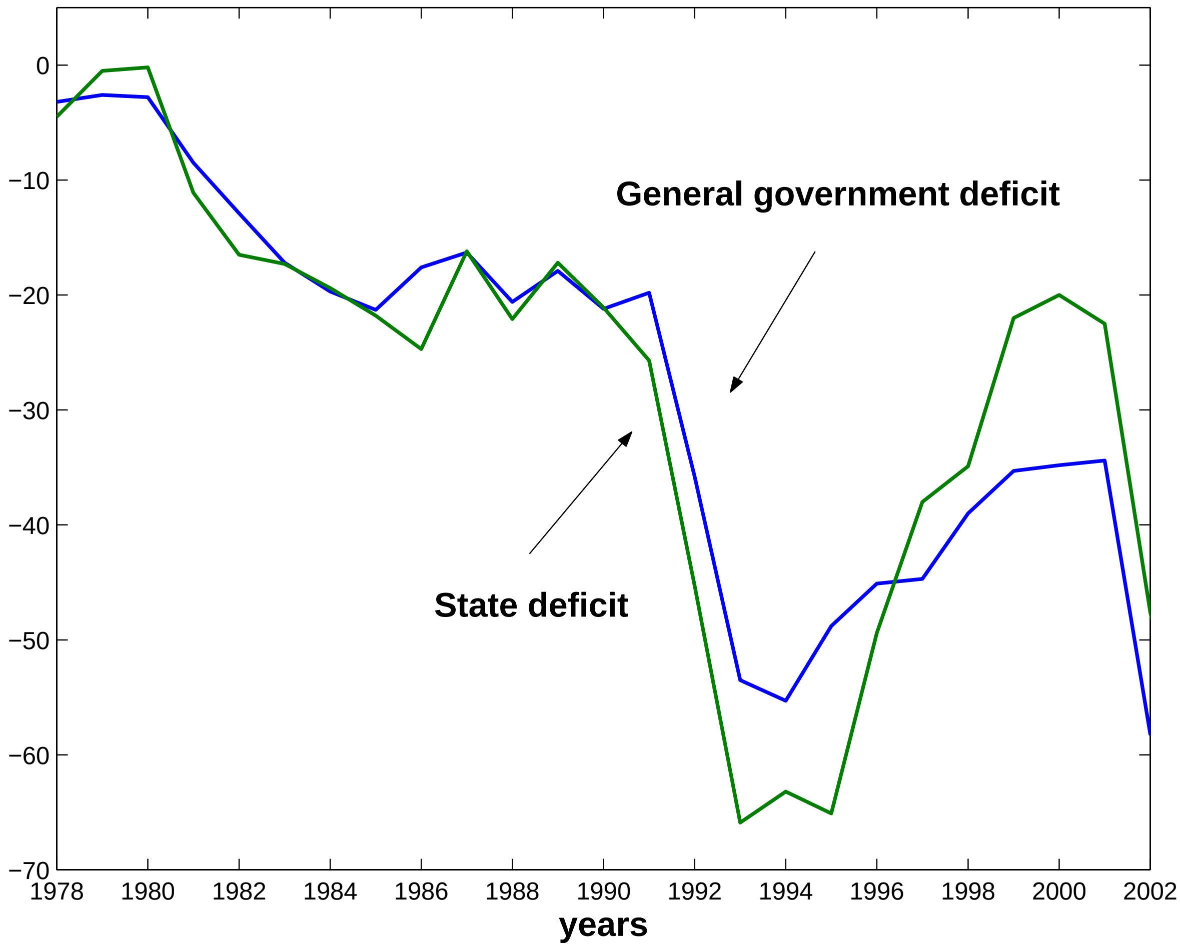

Figure 1: Evolutions of the State and general government deficits. y-axis is in billion of euros

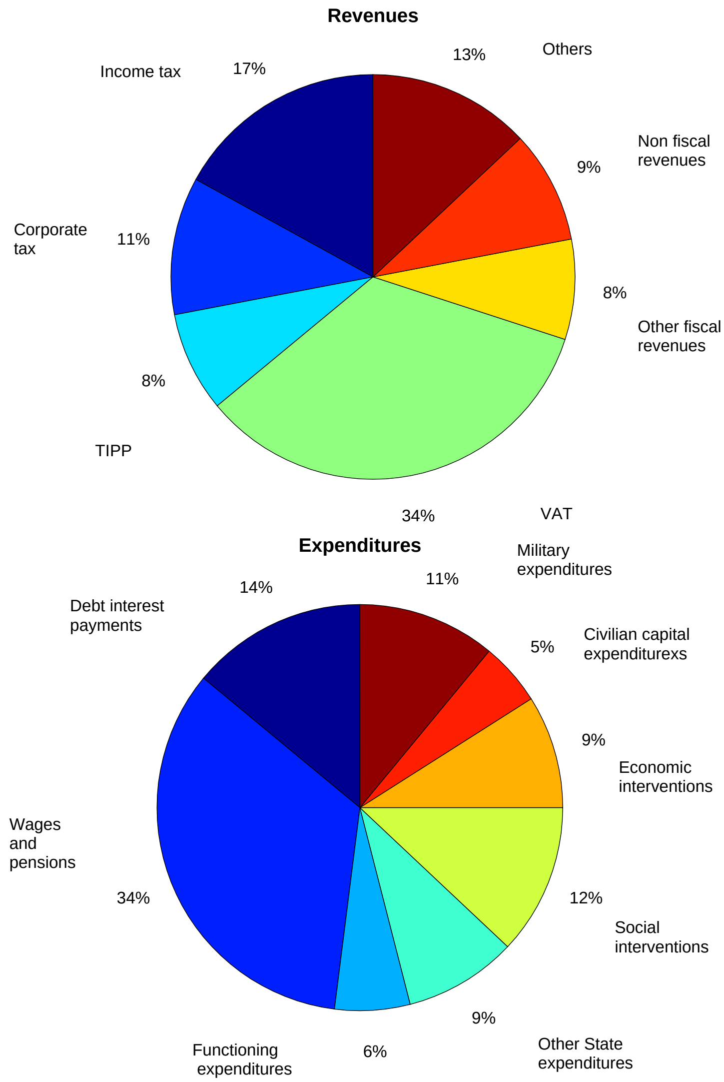

Figure 2: Revenues and expenditures structure.

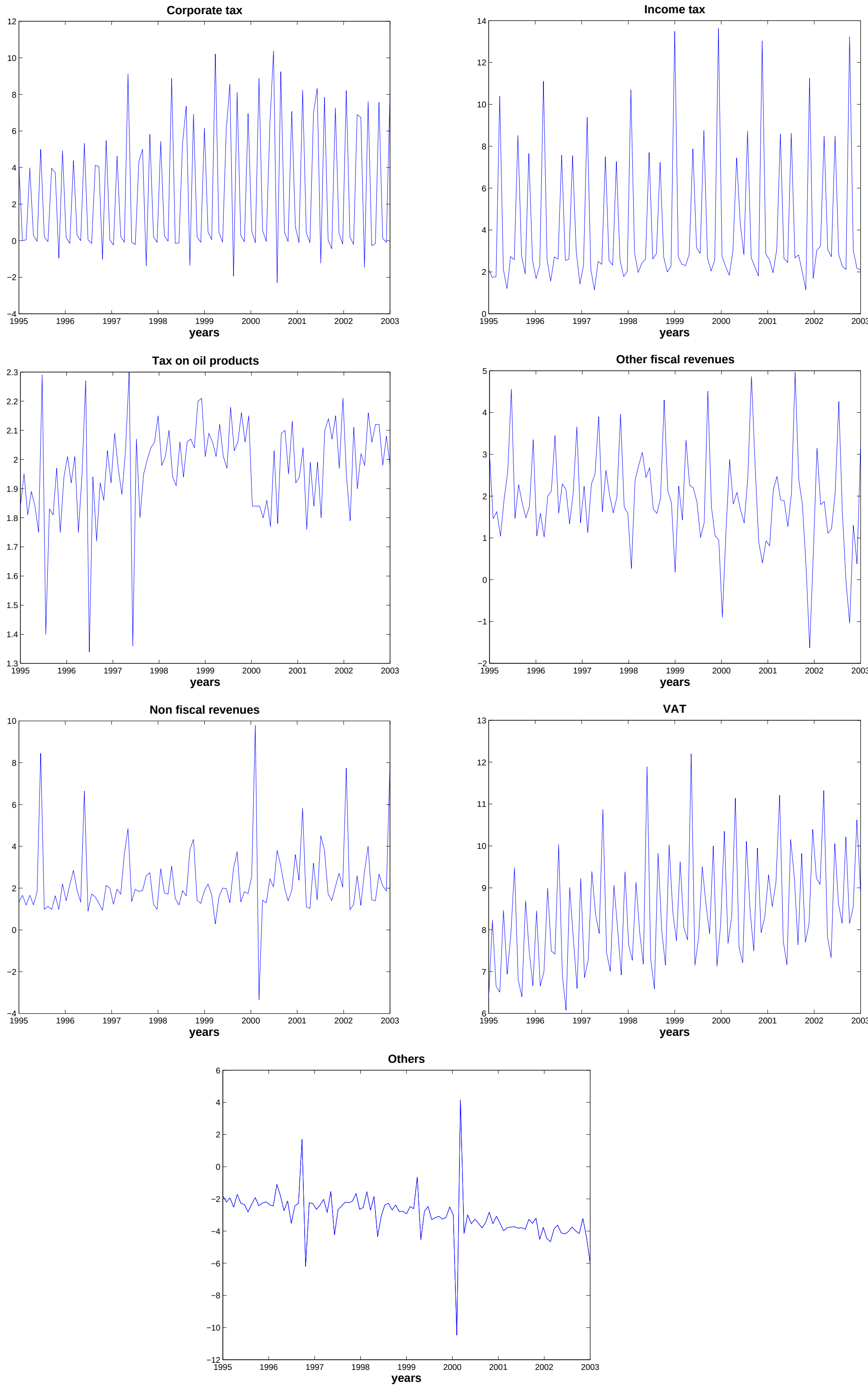

Figure 3: Time series plots of revenue side. y-axis is in billion of euros

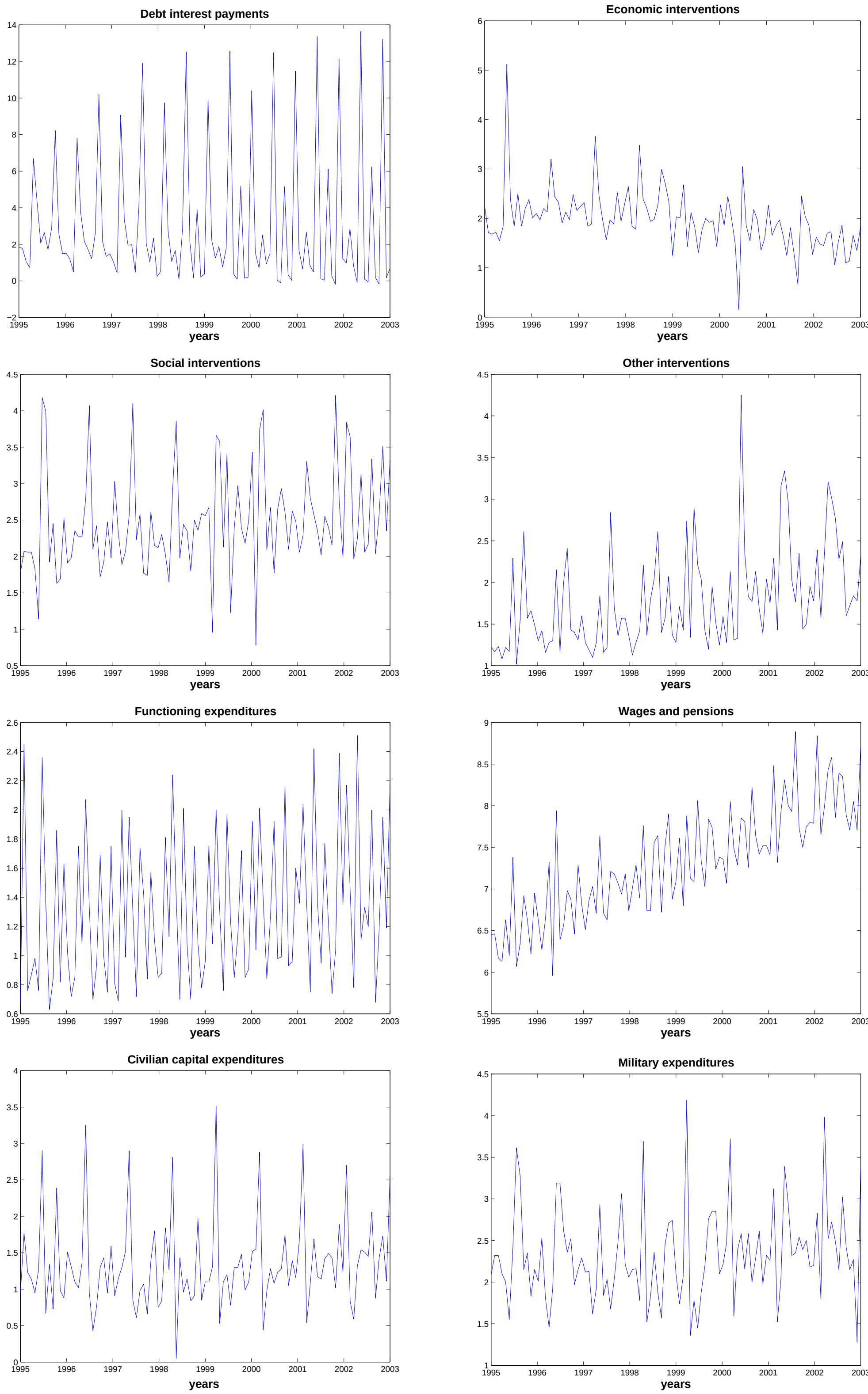

Figure 4: Time series plots of expenditure side. y-axis is in billion of euros

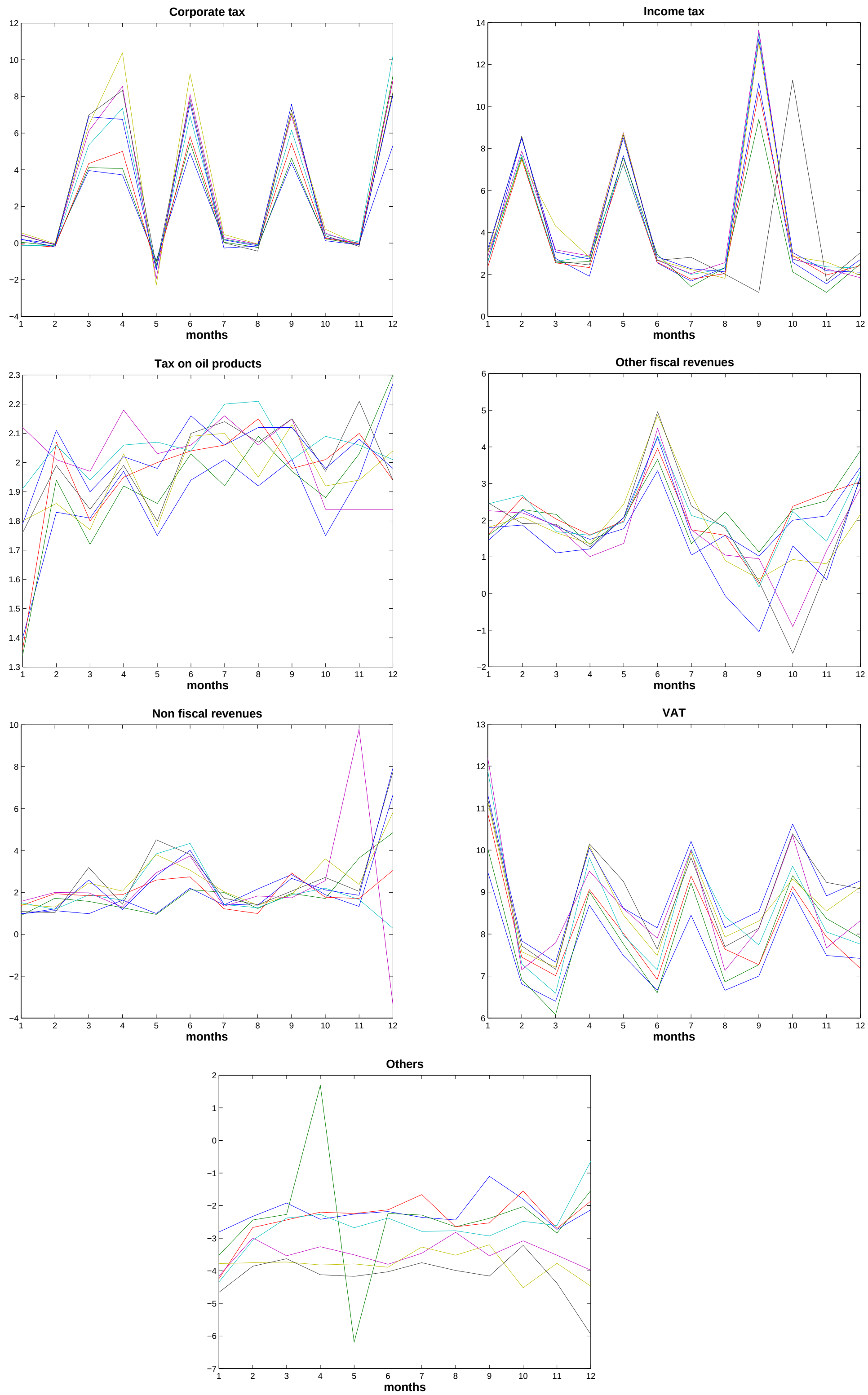

Figure 5: Seasonal patterns on revenue side. y-axis is in billion of euros

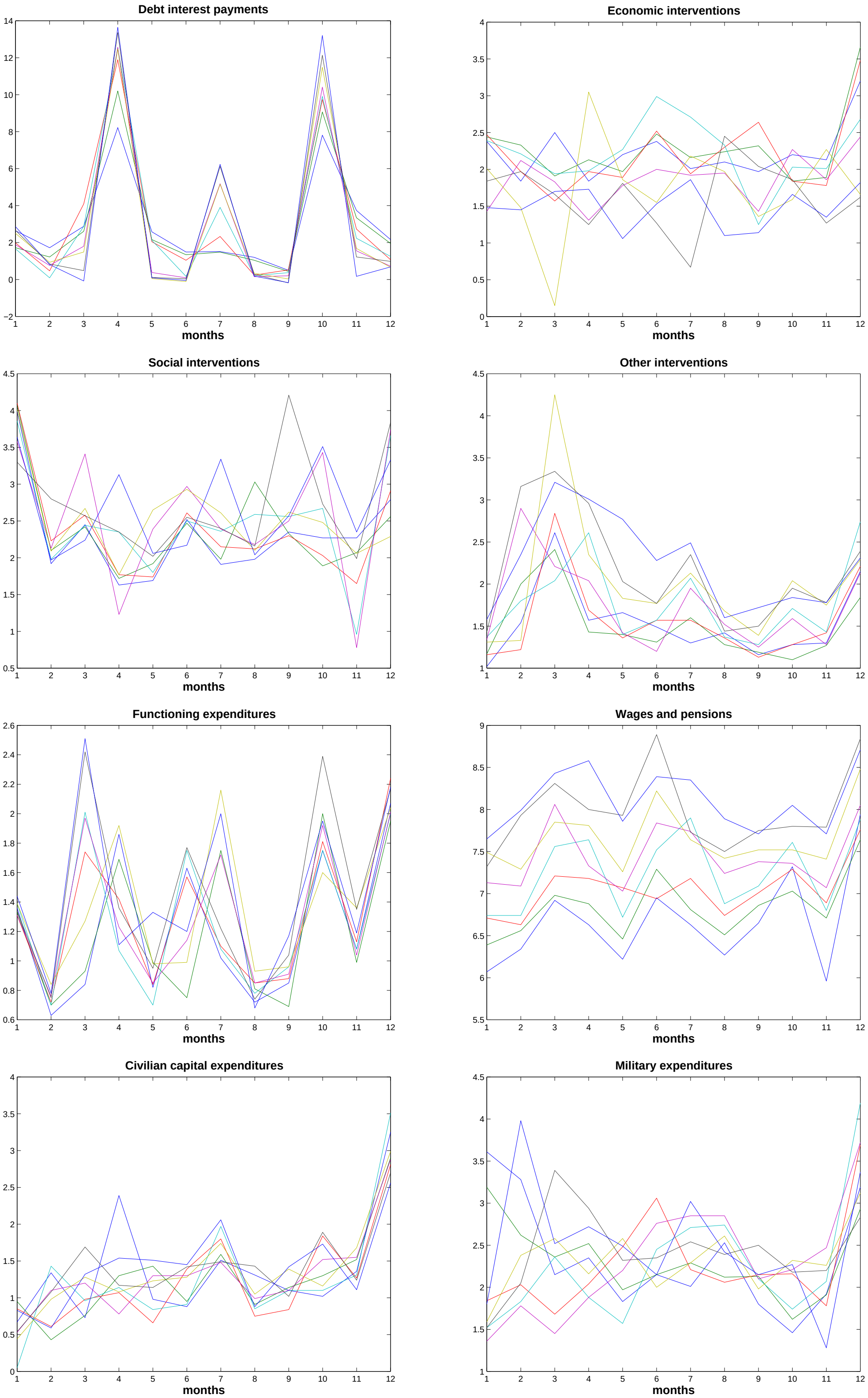

Figure 6: Seasonal patterns on expenditure side. y-axis is in billion of euros

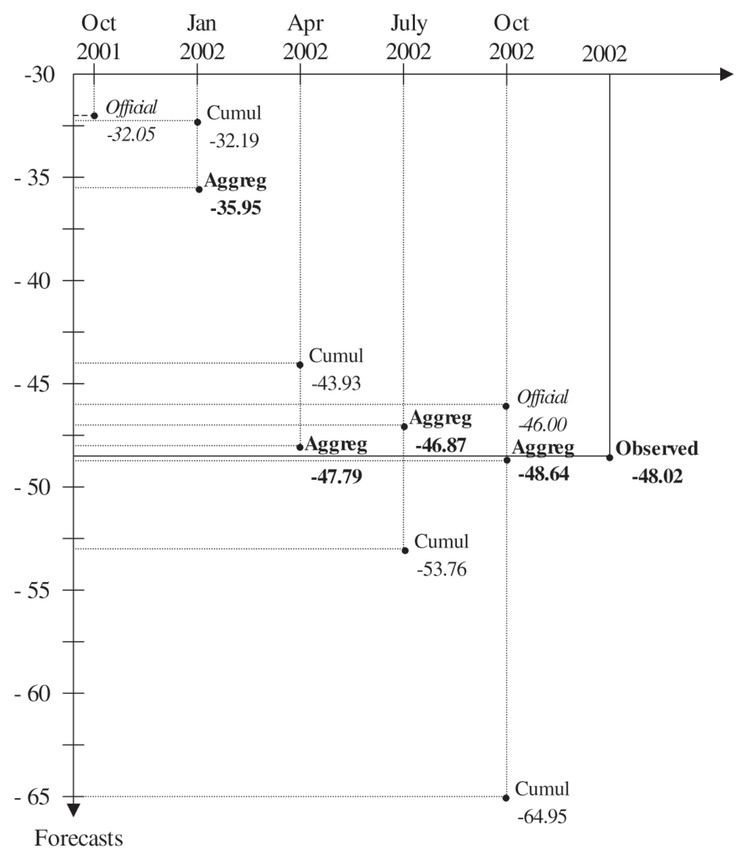

Figure 7: Forecasting State deficit for 2002. y-axis is in billion of euros

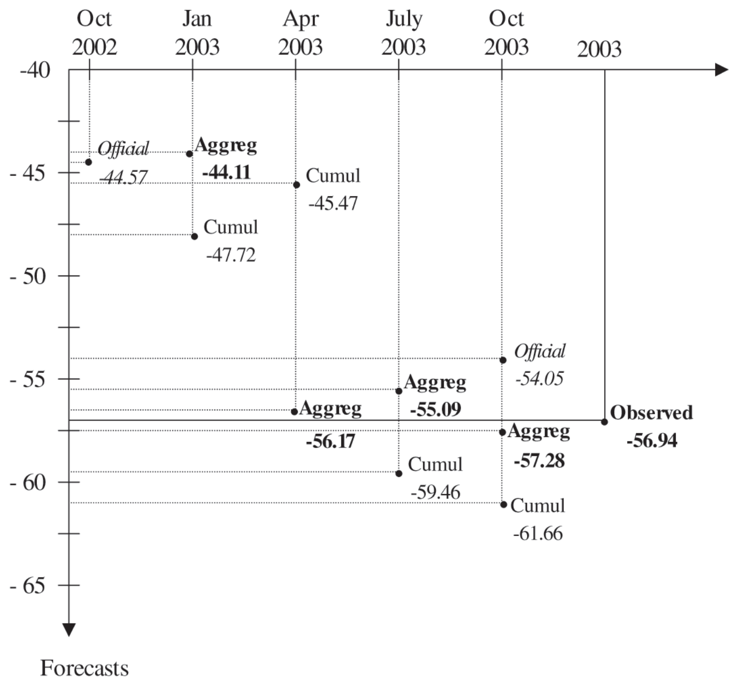

Figure 8: Forecasting State deficit for 2003. y-axis is in billion of euros