HOW RELIABLE ARE THE STATISTICS FOR THE STABILITY AND GROWTH PACT?

2007

Luis Gordo Mora and João Nogueira Martins

Notas Estadísticas
N.º 4

BANCO DE ESPAÑA
Eurosistema

HOW RELIABLE ARE THE STATISTICS FOR THE STABILITY AND GROWTH PACT?

HOW RELIABLE ARE THE STATISTICS FOR THE STABILITY AND GROWTH PACT? (*)

Luis Gordo Mora (**)

BANCO DE ESPAÑA

João Nogueira Martins (***)

EUROPEAN COMMISSION

(*) The views expressed are those of the authors and should not be attributed to the Banco de España or to the European Commission. The authors acknowledge helpful comments by F. Balassone, M. Buti, A. Déruennes, G. Giudice, F. Lequiller, J. Kröger, A. Pina, A. Turrini, S. Zotteri and several other colleagues in the Banco de España and the Commission, as well as editorial assistance by S. Bland. The paper was prepared during an internship of L. Gordo in the Commission, from October to December 2006.

(**) Banco de España – Directorate-General Economics, Statistics and Research. Statistics and Central Balance Sheet Data Office (lgordo@bde.es).

(***) European Commission – Directorate-General for Economic and Financial Affairs (joao.nogueiramartins@ec.europa.eu).

Notas Estadísticas. N.º 4
2007

This series disseminates papers of a statistical nature prepared by the Banco de España. The Notes address methodological aspects such as the scope, sectorisation and recording of financial stocks and transactions; characteristics of securities markets and of interest rates; background notes and others.

The opinions and analyses in the Notas Estadísticas (Statistical Notes Series) are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the INTERNET at the following website: http://www.bde.es

ISSN: 1699-3985 (on line)
Imprenta del Banco de España

Abstract

The aim of this paper is to assess the reliability of the government deficit and debt figures reported to the European Commission by Member States. Reliability is one of the several dimensions of quality in statistics; it refers to the magnitudes of data revisions after the publication of the first outcomes. The measurement of the data reliability and inference about potential future revisions are particularly relevant for fiscal surveillance in the EU since statistical institutes take a long time – usually four years – to provide final data, while the decisions on the Stability and Growth Pact context are taken on the basis of the first estimates available shortly after the end of each year. The paper shows that there are very significant differences in reliability among Member States and indicates the margins of uncertainty in relation to the most recent years' data. It also compares the reliability of deficit and debt figures; checks that the shift from ESA79 to ESA95 did not generally harm the reliability of data and suggests that the size of deficits may have an impact on the way statistical offices revise data.

Keywords: reliability, quality of statistics, Stability and Growth Pact, government deficit

1 Introduction

> *"The fundamental objective of the excessive deficit procedure is to identify, as quickly as possible, a situation of excessive deficit or debt and to put an end to it."*
> (ECOFIN Council Conclusions on statistical governance, 8 November 2005.)

The fundamental objective of the excessive deficit procedure (EDP) – that is, the corrective arm of the Stability and Growth Pact (SGP) – is to identify, and put to an end to, situations of excessive deficit, as quickly as possible.[1] This aim is reflected in the rapidity with which several steps in the procedure take place. Interestingly, the title of the applicable regulation is "on speeding up and clarifying the implementation of the excessive deficit procedure."[2] Although the 2005 reform of the SGP has extended the deadlines for several steps in the procedure, made those deadlines less strict and the whole procedure less tight, the whole procedure still takes place at a frantic pace.[3]

This haste contrasts with the time frame for compilation and publication of budgetary statistics. The first outcomes for a given year are published at the beginning of the subsequent year – but these figures are not final; they include partial information or estimates that need to be revised at a later stage. Data are revised twice a year for at least four years; sometimes more frequently and for longer.

The difference in the time frames of policy decisions and of availability and revision of data means that the first outcomes are those that matter for the procedure and those on the basis of which crucial policy decisions are taken. Therefore, the strategic incentives linked to numerical rules relate first and foremost to the first data transmissions (see for example Buti et al., 2006). The revised data are of better quality than the first outturns but are rarely relevant for the procedure, as they appear too late.

In a number of instances, decisions would have been different if final data had been available earlier.[4] The large revisions in the data of a few countries have harmed not only those countries' reputation, but also the credibility of the whole procedure, and have raised issues of equity among Member States.

---

1. For the economic foundations of fiscal surveillance in the EU, see Brunila, et al. (eds.) (2001) and Buti and Franco (2005) and references therein.

2. Council Regulation (EC) N° 1467/97 (OJ L 209, 2.8.1997, p.6). Regulation as amended by Regulation (EC) N° 1056/2005 (OJ L 174, 7.7.2005, p. 5).

3. Data are reported by Member States in spring and autumn each year. Eurostat has three weeks to scrutinise, validate and publish the data. If the reference values are exceeded, the Commission prepares a report. Within two weeks of the adoption of the report, the Economic and Financial Committee formulates an opinion; within four months of the reporting dates, the Council decides whether the deficit is excessive and adopts recommendations; the country concerned has six months for effective action; immediately after this period, the Council checks whether effective action has been taken; if not, it has two months to adopt a notice. A decision to impose financial sanctions can be taken four months after the notice and 16 months after data have been reported. For more on the steps of the procedure and on the SGP reform, see European Commission (2005, in particular section II.1 thereof).

4. A few examples: the excessive deficit procedure for Finland was abrogated in June 1997 on the basis of a deficit ratio for 1996, reported on 1 March 1997, of 2.6% of GDP; that ratio was then successively revised to 3.1%, 3.3% and even 3.5% of GDP. In March 2002, Portugal reported a deficit ratio of 2.2% of GDP for 2001; after six months the deficit ratio was revised to 4.1% of GDP; as a result, the ensuing procedure started half a year later than should have been the case; in March 2004, two years after the first transmission, the deficit ratio was further revised upwards to 4.4% of GDP. In Greece, there were several very large revisions in the deficit and debt ratios; the largest revision concerned 2001: Greece reported a small government surplus of 0.1% of GDP in March 2002 which was then successively revised to deficits of 1.4%, 1.5%, 3.7% and 6.1% of GDP. In spring 2005, the Commission initiated a procedure against Italy, as it appeared that the 2004 deficit was at 3.0%; in fact the 2004 deficit ratio was revised to 3.4% and it appeared that the 2001 and 2003 ratios were also above 3%.

Are there ways of improving this situation? Clearly, the solution is not to stop the SGP-related procedures and wait until final data are available. In many cases, when final figures are published, the budgetary difficulties that motivated a procedure have already been solved or, in a less benign configuration – which would be likely in case if no policy action has been taken, the situation has kept deteriorating, and the correction will be more painful.

Over the last years, there have been some measures at the EU level and in specific Member States intended to upgrade the quality of the budgetary statistics.[5] While the very large revisions that took place in a few countries in the past were exceptional and will hopefully not be repeated, it must be acknowledged that data revisions are inevitable. Although technological progress in information technology speeds up data collection, the complexity of the accounting rules and the number of transactions to record (millions of transactions by thousands of government units in each country) mean that the production of final statistics remains a long and painstaking task.

One possible way of improving things would be to explicitly acknowledge, in each step of the EDP, that the budgetary data are subject to revisions, and that the reliability of data increases with time and differs from country to country. Though we are unable to say how the first outcome published by a given country will be revised, experience can help us infer the uncertainty surrounding the deficit and debt figures reported by Member States. This may help in identifying situations of risk of excessive deficit. In some cases, it may make sense to postpone the abrogation of a procedure when the deficit ratio has fallen below, but remains close to, 3% of GDP, and there is a significant risk that the deficit figures will have to be revised to above the reference value at a later stage, in particular if some other indicators suggest that an upward revision is likely.[6] At least the policy-makers should be made aware of, and the relevant acts should explicitly refer to, the fact that the data on which crucial decisions are taken are subject to revision.[7]

The aim of this paper is to assess the reliability of the government deficit and debt figures reported to the European Commission by Member States.[8] As far as we know, no other paper has attempted to do so in a systematic manner. The paper is organised as follows. In the next section, we discuss some conceptual issues about quality in statistics and better delimit the scope of the analysis; in particular, we identify reliability as one of the several dimensions of quality in statistics. Section 3 details the available data and the procedure for calculating the relevant indicators. Sections 4 and 5 compare the reliability of figures reported by the different EU Member States. Section 6 compares the revisions in deficit and debt figures. Section 7 elaborates on the evolution of reliability over time, while section 8 provides a preliminary examination on the relationship between size of revisions and size of deficits. Section 9 summarises the main findings and concludes.

---

5. See European Commission (2005, in particular box II.2; and 2006, in particular section II.4.4 thereof).

6. On the usage of multiple indicators to countercheck the official deficit and debt figures, see Balassone et al. (2006a).

7. Another way to accelerate the compilation of statistics could be to simplify the accounting rules. However, this option does not seem to be on the cards and this paper does not dwell on this specific issue. Note, nonetheless, that Tanzi (2006) argues that "the concept that the statisticians have chosen (the accrual concept), though superior to [a simpler] cash concept in some ways, suffers from the shortcoming that it is not easy to calculate and does not provide policy-makers with a gauge of the fiscal situation at the time when it is most needed – during and not after, or much after, the fiscal year". Balassone et al. (2003) recall that the overall statistical framework and its complexity was not designed to support the enforcement of rules and even suggest redefining the reference deficit as the change in nominal debt, with benefits in timeliness and transparency. Buti (2006) also pleads for simple rules as a way of increasing the reliability of data.

8. We are not aware of any similar paper for the US. The comparison between revisions in the US government finance statistics and the EU Member States is a topic for further research.

2 Quality of Statistics: A Multidimensional Concept

The quality of economic statistics is crucial for an adequate understanding of the economic situation and for effective policymaking. Poor-quality statistics may lead to poor economic analysis, mistaken conclusions about the behaviour of economic agents and inappropriate policy decisions. The quality of the budgetary statistics of Member States is particularly important given that these statistics are the foundation of a budgetary surveillance framework in a multi-country context (European Commission, 2003, in particular section II.4 thereof).

The quality of statistics can be generically defined as fitness for use (OECD, 2003). This encompasses all aspects of how well statistics meet users' needs and their expectations about the information content of the disseminated data (Damia and Picón Aguilar (2006), quoting the Encyclopaedia of Statistical Sciences). The quality of statistics is a multidimensional concept. The data quality assessment framework (DQAF) for government finance statistics of the IMF (2003) identifies the following dimensions of quality: (i) integrity, (ii) methodological soundness, (iii) accuracy and reliability, (iv) serviceability and (v) accessibility, as well as a number of (vi) prerequisites of quality. These dimensions of quality can be broken down into a number of quality elements, notably relevance, transparency, timeliness, punctuality, consistency, completeness, comparability across time and space, interpretability, credibility, plausibility and cost-efficiency. The UN, OECD and Eurostat identify slightly different dimensions of statistical quality, or define them in slightly different ways, but differences are more semantic than substance.[9]

Statistics are accurate if they sufficiently portray reality. Accuracy is thus related to the discrepancy between statistics and the unknown "true" figures they are intended to measure. This discrepancy can be separated into two components. The first is the difference between final figures and reality. The measurement of this discrepancy is very complex, since the "true" figures are not observable. The second is reliability. This is the difference between the first published data and the final figures. (Bier and Ahnert, 2001, and Carson and Laliberté, 2002). Therefore, reliability (or stability) of data refers to the closeness of the initial estimated values – which are those which are more relevant for policymaking – to the subsequently revised values.

The dimension of statistical quality that is discussed in the paper is reliability. However, several dimensions of statistical quality, other than reliability, are of the utmost relevance for fiscal surveillance. There are interesting parallels between the dimensions of statistical quality and the ideal qualities of fiscal rules as presented in the academic literature: namely that they should be well-defined, transparent, simple, flexible, adequate to final goal, enforceable, consistent, and linked with structural reform (Kopits and Symansky, 1998). Fiscal rules will not be enforceable if fiscal statistics are not timely and reliable; the transparency of the fiscal numerical rules is directly connected with the transparency of the available data; simple rules (e.g. those that refer to a single numerical indicator with a small or no margin for judgement) will be adequate to the final goal only if the numerical indicator is very accurate and relevant; more flexible fiscal rules which allow a wider margin of judgement will be less dependent on the accuracy of the numerical indicators (Buti, 2006).

---

9. See e.g. De Vries (2002), Eurostat (2003), IMF (2003), OECD (2003), Saeboe (2003), Giovannini (2004). For a comparison between the IMF's data quality assessment framework and Eurostat's quality definition, see Laliberté et al. (2004) and Laliberté and Defays (2006).

Moreover, the dimensions of statistical quality are not independent from each other. There are trade-offs between several dimensions, for example between timeliness and reliability (Bier and Ahnert, 2001); integrity, reliability and accuracy; completeness and comparability; methodological soundness and comparability across time; etc.

Therefore, any conclusion on the reliability of data of any country cannot be understood as an overall assessment of the quality of their data, nor on the fitness of these data for use in fiscal surveillance. The reliability of government data is a necessary but by no means sufficient condition for the fitness of data for use in fiscal surveillance.

3 Data Available and Reliability Indicators

The EU Member States compile and publish government deficit and debt figures at their own pace. For example, in Germany, the deficit and debt figures are usually published very shortly after the end of the respective year. Deficit data for 2005 were published by the German Statistical Office on 12 January 2006; a first revision, together with detailed revenue and expenditure accounts, was released on 22 February 2006.[10] Other EU countries publish their data somewhat later.

However, all EU Member States are under an obligation to report their deficit and debt figures to the Commission at precise dates in spring and autumn each year. Such figures – together with the Commission services' forecasts, which are published shortly after the fiscal reporting – constitute the raw material on the basis of which fiscal surveillance procedures are put in motion. Each data transmission covers the previous four years. After four years, or eight notifications, it is assumed that data are final; they then drop from the data transmission. In other words, data for year t are reported for the first time in spring t+1 and revised until autumn t+4: 42 months after the transmission of the first outcome, or 45 months after the end of the year to which they refer. These extraordinarily long lags for the publication of final data illustrate just how demanding the task of compiling government deficit and debt figures is. Some countries keep revising their data after four years; however, for the purpose of this paper, we will assume that data of all countries are final after four years.

The available sample consists of data reported to the Commission by Member States between March 1994 and October 2006.[11] Each data transmission includes data with different revision histories. In spring t, data for t-1 are transmitted for the first time, data for t-2 are transmitted for the third time, data for t-3 are transmitted for the fifth time and data for t-4 are transmitted for the seventh time. The second, fourth, sixth and eight transmissions take place in autumn. To simplify, we will say that each data transmission (also called a data vintage) contains data of different ages: ages 1, 3, 5 and 7 in spring, and 2, 4, 6 and 8 in autumn.

For each country, there are 104 observations, that is, 13 observations of each age.[12] The analysis concerns 14 countries: Belgium, Denmark, Germany, Spain, France, Ireland, Italy, Luxembourg, Netherlands, Austria, Portugal, Finland, Sweden and the United Kingdom. We have excluded Greece from most of the analysis, given that the very large revisions of a structural nature, which took place between autumn 2002 and autumn 2006, cannot be used to infer anything about potential future revisions. Because of data availability issues, the assessment of the reliability of reported deficit figures does not concern the ten new Member States, which have reported their data under the same conditions as other Member States since spring 2004 only.

Data of age 2 are expected to be more reliable – i.e. closer to final data – than data of age 1, data of age 3 are expected to be more reliable than data of age 2, and so on. As a result, for each of the 14 countries considered, there are 8 different samples of 13

10. See Destatis Press Release of 22 February 2006.

11. Deficit and debt data reported by Member States since 1994 are available from the authors upon request.

12. The number of observations for Austria, Sweden and Finland, which have reported their data to the Commission since their accession to the EU in 1995, is 96.

observations each. These are short samples for statistical inference purposes; as a result, reliability indicators need to be interpreted carefully and prudence is necessary in the conclusions.

The most relevant statistics for the implementation of the Stability and Growth Pact are the ratios-to-GDP of the government deficit and debt. Revisions in the deficit and debt ratios can occur because of revisions in the numerators and denominators, that is, because of revisions in the deficit and debt figures, or revisions in the GDP levels. This paper is concerned about the revisions in the deficit and debt levels only. We disregard the fact that the nominal GDP figures are also subject to revisions.[13] However, since there is a need to standardise data, to make them comparable across time and across countries, our calculations are as a percentage of the latest available GDP series.

Table 1 summarises the available sample of deficit revisions. The number of revisions falling in each of five brackets (large and medium-size revisions that reduce the deficit, small revisions, and large and medium-size revisions that increase the deficit) is indicated. In relation to the central bracket corresponding to small revisions, the table also reports on the deficit-increasing revisions and on very small revisions. As expected, the number of small revisions (revisions in the deficit ratios between -0.2 and 0.2% of GDP), and very small revisions (revisions between -0.05% and 0.05% of GDP) tends to increase with the data ages, while the number of very large revisions (the first and fifth column corresponding to revisions in the deficit ratios below -0.5% and above 0.5% of GDP) falls. Moreover, the table suggests normality in the revisions (see Annex I.3).

Table 1: Sample available – Revisions in government deficit between two consecutive data transmissions

	Size of revisions in % of GDP							
	x<-0.5%	-0.5%<x<-0.2%	-0.2%<x<0.2%			0.2%<x<0.5%	0.5%<x	TOTAL
			Total	Of which: 0.0%<x<0.2%	Of which: -0.05%<x<0.05%			
1st	7	16	123	62	58	25	11	**182**
2nd	4	20	125	55	59	15	4	**168**
3rd	6	13	147	55	85	13	3	**182**
4th	6	17	131	53	82	11	3	**168**
5th	2	7	160	51	110	11	2	**182**
6th	2	16	144	54	104	2	4	**168**
7th	2	5	167	48	129	6	2	**182**
TOTAL	**29**	**94**	**997**	**378**	**627**	**83**	**29**	**1232**

Note: The table summarises the revisions in government deficit data in the sample. Revisions are classified according to the number of revisions (lines) and their size (columns). A first revision for a given year corresponds to the difference between the first and the second data transmissions, i.e., data for year t reported in spring t+1 and autumn t+1. Second revisions correspond to the differences between the second and the third data transmissions, i.e. data for year t reported in autumn t+1 and spring t+2, and so on. Positive revisions (+) mean increases in deficits or reductions in surpluses; negative revisions (–) mean reductions in deficits or increases in surpluses. As expected, the number of small revisions (between -0.2 and 0.2% of GDP) and very small revisions (between -0.05% and 0.05% of GDP) increases with the data age, while the number of very large revisions (revisions in the deficit ratios below -0.5% and above 0.5% of GDP) falls.

Not all data revisions have the same nature. Data are typically revised because new information on government transactions becomes available, errors or inconsistencies are identified and corrected, or as a result of the continuous efforts to better comply with the

13. See Annex I. On GDP revisions, see also Aruoba (2005) and Di Fonzo (2005).

intricacies of the accounting rules: these are *routine* revisions. We may conjecture that the size of the routine revisions differs among countries because of the size of the country, the institutional arrangements in place, including the degree of decentralisation of government, the technical arrangements put in place to collect data from each government unit, the habits and obligations in the production of source data by each of the government units, the similarity between the accounting standards in use in each of government unit and the national accounting rules, the available resources and technologies to compile the first outcome and to revise data, as well as the complexity of the transactions carried out by each country, etc. There are also *structural* revisions of data series, when the accounting rules are significantly revised (for example the shift of ESA79 to ESA95 at the beginning of 2000, or clarifications to accounting rules published by Eurostat[14]), major systemic changes in the data compilation, or when exceptionally large errors in the accounts are identified and corrected.[15]

In this paper, we are interested in calculating indicators of reliability, and inferring the size of future revisions. Because of their nature, nothing can be inferred about future revisions on the basis of the structural revisions in the past. The fact that ESA79 was replaced with ESA95 as the accounting standard for the compilation of government accounts in spring 2000, and that such a methodological change led to revisions in government accounts, does not suggest similar revisions will take place in the coming years, when they will take place, or what will be the consequences for the time series of another large methodological change of the similar nature. In principle, we are interested in the routine revisions only.[16]

Therefore, structural revisions should be identified and removed from the sample. However, given the difficulty in identifying structural revisions and of distinguishing them from routine revisions, we have abstained from attempting doing so in a systematic way.[17] The only structural revisions that we have specifically identified and removed from the sample are those corresponding to the shift from the ESA79 to ESA95 at the beginning of 2000, which was relevant for all Member States, and the breaks in the Danish series in connection with the sectoral classification of pension schemes (see Annex II.1). This implies that, for a number of countries that have undergone structural revisions in their statistical system over the last fifteen years or so, we may be underestimating the reliability of their data.[18]

---

14. Since 1996, Eurostat has regularly published clarifications on the implementation of ESA. These clarifications are compiled in a “Manual on government deficit and debt”. The manual, together with updates, is available at: http://epp.eurostat.ec.europa.eu/portal/page?_pageid=2373,47631312,2373_
58674363&_dad=portal&_schema=PORTAL.

15. Aruoba (2005) uses the terms ‘informative’ and ‘uninformative’ revisions to refer to what we call ‘routine’ and ‘structural’ revisions.

16. For example, the very large revisions in the Greek government accounts in 2002 and from 2004 to 2006 were of a structural and systemic nature and cannot be taken as an indication that similar revisions will occur in the future.

17. It should be noted that structural and routine revisions are not distinguishable by their size, but by their nature, and that structural and routine revisions may either add up or offset each other. The revision in the Finnish figures quoted in footnote 4 was clearly a routine revision, though it amounted to almost 1% of GDP. Most revisions because of the structural shift from ESA79 to ESA 95 were much smaller in size.

18. An illustration of the effect of removing structural revisions other than the shift from ESA79 to ESA95 for Portugal and Sweden – two of the countries most affected by structural breaks – features in Annex II.2.

4 Comparing the Reliability of Deficit Data among Different Countries

Studying how deficit figures have been revised since 1994 allows Member States to be compared and ranked. Tables 2 to 4 show, for each Member State, data on standard deviations in revisions, average of revisions and average of revisions in absolute values between each data age (1 to 7) and final data (age 8). The most relevant indicators correspond to the revisions from the first transmission (of spring t+1) to final data (of autumn t+4); however, the tables show indicators for each data age.[19]

The standard deviation of the deficit revisions is shown in table 2. It indicates the precision of the average revision as an estimator of future revisions. According to this indicator, France, Germany and the United Kingdom are the countries that have reported the most reliable deficit figures. For France and Germany, the standard deviation of revisions from the first transmission to final data has been below 0.3% of GDP, while the UK data are only slightly above that mark (see column 1). The French figures remain the most reliable until age 5, when the data of almost all countries are already very close to final. It is remarkable that the

Table 2: Member States ranked by the standard deviation of revisions in deficit data (% of GDP)

	Revisions cumulated from						
	1st transmission to final	**2nd transmission to final**	**3rd transmission to final**	**4th transmission to final**	**5th transmission to final**	**6th transmission to final**	**7th transmission to final**
1	FR(0.23%)	FR(0.21%)	FR(0.16%)	FR(0.14%)	FR(0.11%)	NL(0.06%)	SE(0.00%)
2	DE(0.29%)	UK(0.23%)	UK(0.18%)	UK(0.15%)	UK(0.12%)	FR(0.08%)	FI(0.04%)
3	UK(0.31%)	DE(0.23%)	FI(0.21%)	DE(0.18%)	DE(0.12%)	AT(0.08%)	FR(0.05%)
4	BE(0.37%)	NL(0.29%)	DE(0.23%)	NL(0.18%)	AT(0.13%)	UK(0.09%)	AT(0.05%)
5	IE(0.39%)	FI(0.30%)	NL(0.27%)	FI(0.19%)	FI(0.13%)	DE(0.09%)	NL(0.06%)
6	FI(0.40%)	BE(0.34%)	AT(0.27%)	AT(0.22%)	NL(0.18%)	FI(0.13%)	UK(0.07%)
7	ES(0.42%)	IE(0.35%)	IE(0.30%)	ES(0.26%)	IT(0.23%)	LU(0.18%)	DE(0.08%)
8	AT(0.45%)	AT(0.36%)	BE(0.33%)	IE(0.26%)	IE(0.23%)	ES(0.18%)	PT(0.11%)
9	NL(0.51%)	ES(0.36%)	ES(0.34%)	BE(0.29%)	ES(0.25%)	IE(0.19%)	LU(0.11%)
10	IT(0.53%)	IT(0.48%)	IT(0.43%)	PT(0.32%)	LU(0.26%)	BE(0.20%)	IT(0.12%)
11	DK(0.74%)	PT(0.54%)	PT(0.47%)	IT(0.33%)	PT(0.28%)	IT(0.21%)	BE(0.17%)
12	PT(0.78%)	DK(0.67%)	DK(0.61%)	LU(0.41%)	BE(0.29%)	PT(0.25%)	ES(0.17%)
13	LU(0.92%)	LU(0.71%)	LU(0.62%)	DK(0.60%)	SE(0.53%)	DK(0.48%)	IE(0.18%)
14	SE(1.01%)	SE(1.00%)	SE(0.88%)	SE(0.87%)	DK(0.56%)	SE(0.53%)	DK(0.47%)

Note: The table shows the standard deviations of revisions between any given data transmission and final data. The first column corresponds to the revisions between the first outturn for year t reported in spring t+1 and final data for the same year as transmitted in autumn t+4. A higher value in standard deviation of the revisions means more dispersion in deficit data for a given year over different data transmissions and a worse position of the country in terms of reliability. The reduction in the standard deviations for each country when advancing in data transmissions (from 1st to 2nd and so on) shows the pace at which data progress to final.

19. Annex I provides some further explanations about the process of calculating these indicators and some underlying issues.

French figures of age 1 are more reliable than the data of many other countries of more advanced ages. For example, it is only at the 5th transmission that the Italian figures reach the same reliability of French figures of age 1. Denmark, Sweden and Luxembourg are the countries with the largest dispersion of the revisions from age 1 to final data. In particular, the Danish and Swedish data[20] remain much more susceptible to revisions compared to other Member States until ages 6 and 7, as revisions in their data have tended to be published relatively late.[21]

As far as the average revisions are concerned (Table 3), the two countries with the largest deficit-reducing (or surplus-increasing) revisions from data of age 1 to final (column 1) are Luxembourg and Belgium. The two countries with largest average of deficit-increasing revisions are Portugal and Italy.[22] It is worth noting that Portugal and Italy appear among the countries with the largest deficit-increasing revisions until age 4 (in autumn t+2). Note, however, that given the large standard deviation reported above, the average revisions in all countries are not significantly different from zero at the usual significance levels. It is remarkable that, while the Swedish figures have been subject to large revisions, as measured by the standard deviations (Table 2), Sweden has been among the countries with the smallest average revisions, that is, its revisions have cancelled each other.

Table 3: Member States ranked by the average of revisions in deficit data (% of GDP)

	Revisions cumulated from						
	1st transmission to final	2nd transmission to final	3rd transmission to final	4th transmission to final	5th transmission to final	6th transmission to final	7th transmission to final
1	LU(0.75%)	LU(0.14%)	SE(0.15%)	SE(0.11%)	SE(0.19%)	SE(0.19%)	IE(0.06%)
2	BE(0.27%)	BE(0.10%)	LU(0.09%)	FI(0.08%)	IE(0.05%)	IE(0.08%)	BE(0.05%)
3	IE(0.09%)	FI(0.08%)	BE(0.09%)	IE(0.05%)	DE(0.02%)	DE(0.02%)	DE(0.02%)
4	DK(0.05%)	IE(0.05%)	FI(0.07%)	DK(0.03%)	FI(0.02%)	BE(0.01%)	FI(0.01%)
5	FI(-0.01%)	SE(-0.01%)	IE(0.04%)	FR(0.02%)	ES(0.02%)	FI(0.01%)	IT(0.01%)
6	AT(-0.02%)	FR(-0.02%)	FR(0.01%)	BE(0.01%)	BE(0.01%)	FR(-0.01%)	ES(0.00%)
7	SE(-0.03%)	DK(-0.05%)	DK(-0.02%)	DE(-0.02%)	DK(0.01%)	NL(-0.02%)	FR(0.00%)
8	FR(-0.03%)	UK(-0.05%)	UK(-0.04%)	ES(-0.02%)	FR(-0.01%)	PT(-0.03%)	SE(0.00%)
9	DE(-0.03%)	DE(-0.09%)	DE(-0.07%)	LU(-0.04%)	IT(-0.04%)	ES(-0.03%)	DK(-0.01%)
10	UK(-0.09%)	ES(-0.10%)	ES(-0.11%)	NL(-0.04%)	PT(-0.05%)	AT(-0.03%)	NL(-0.02%)
11	ES(-0.16%)	AT(-0.10%)	AT(-0.11%)	UK(-0.04%)	AT(-0.05%)	IT(-0.05%)	UK(-0.03%)
12	NL(-0.17%)	NL(-0.14%)	NL(-0.12%)	PT(-0.08%)	UK(-0.06%)	UK(-0.05%)	PT(-0.03%)
13	IT(-0.40%)	PT(-0.32%)	PT(-0.22%)	AT(-0.13%)	LU(-0.06%)	DK(-0.06%)	AT(-0.03%)
14	PT(-0.40%)	IT(-0.32%)	IT(-0.23%)	IT(-0.13%)	NL(-0.06%)	LU(-0.08%)	LU(-0.05%)

Note: The table shows the average revisions between any given data transmission and final data. The first column corresponds to the revisions between the first outturn for year t reported in spring t+1 and final data for the same year as transmitted in autumn t+4. A negative (positive) value means an expected increase (reduction) in the deficit.

20. It should be noted that at the 7th transmission, Sweden jumps from the last to the first place in the ranking, when its standard deviation becomes zero. This means that in spite of relatively large errors, the revision process in Sweden stops somewhat earlier than in other Member States.

21. As explained above, Greece was not considered in most of the analysis, and data are not systematically reported in tables and charts. The standard deviation of revisions in deficit data for Greece from the 1st transmission to final data is 1.92% of GDP.

22. The average of revisions in deficit data in Greece from the 1st transmission to final data is -1.60% of GDP.

Table 4 shows the average of deficit data revisions in absolute value for all countries; it complements the results shown in Table 3. Given that Table 4 treats deficit-increasing and deficit-decreasing revisions identically and avoids revisions offsetting each other, the information in Table 4 is quite similar to Table 1, with this indicator showing the French, German and British figures to be the most reliable, and Luxembourg, Sweden and Denmark reporting data subject to the largest revisions in absolute value.

Table 4: Member States ranked by the average of revisions in deficit data in absolute values (% of GDP)

	Revisions cumulated from						
	1st transmission to final	2nd transmission to final	3rd transmission to final	4th transmission to final	5th transmission to final	6th transmission to final	7th transmission to final
1	FR(0.33%)	FR(0.28%)	FR(0.18%)	FR(0.15%)	FI(0.09%)	NL(0.03%)	SE(0.00%)
2	DE(0.44%)	DE(0.31%)	FI(0.23%)	NL(0.16%)	FR(0.10%)	FR(0.06%)	FI(0.01%)
3	UK(0.52%)	UK(0.35%)	UK(0.24%)	UK(0.16%)	DE(0.11%)	DE(0.06%)	NL(0.02%)
4	ES(0.56%)	NL(0.35%)	DE(0.29%)	FI(0.17%)	UK(0.12%)	AT(0.07%)	FR(0.03%)
5	FI(0.60%)	FI(0.40%)	NL(0.30%)	DE(0.18%)	NL(0.14%)	UK(0.07%)	AT(0.03%)
6	NL(0.66%)	ES(0.43%)	ES(0.37%)	ES(0.21%)	AT(0.14%)	FI(0.07%)	DE(0.04%)
7	BE(0.67%)	BE(0.48%)	AT(0.37%)	AT(0.26%)	ES(0.16%)	ES(0.11%)	IT(0.04%)
8	IE(0.69%)	AT(0.50%)	IE(0.43%)	BE(0.34%)	IT(0.19%)	IE(0.14%)	UK(0.05%)
9	IT(0.70%)	IE(0.56%)	IT(0.45%)	IE(0.34%)	IE(0.25%)	IT(0.15%)	PT(0.05%)
10	AT(0.71%)	IT(0.62%)	BE(0.46%)	IT(0.35%)	SE(0.25%)	BE(0.16%)	ES(0.08%)
11	PT(0.93%)	PT(0.67%)	PT(0.51%)	PT(0.37%)	PT(0.27%)	LU(0.18%)	LU(0.09%)
12	DK(1.09%)	DK(0.84%)	DK(0.61%)	DK(0.55%)	BE(0.31%)	PT(0.19%)	IE(0.10%)
13	SE(1.10%)	SE(1.01%)	SE(0.70%)	LU(0.57%)	LU(0.32%)	SE(0.25%)	BE(0.12%)
14	LU(1.86%)	LU(1.20%)	LU(0.92%)	SE(0.65%)	DK(0.43%)	DK(0.31%)	DK(0.22%)

Note: The table shows the average revisions in absolute values between any given data transmission and final data. The first column corresponds to the absolute revisions between the first outturn for year t reported in spring t+1 and final data for the same year as transmitted in autumn t+4. A higher value in this indicator means a worse position on reliability.

Data in Tables 2 and 4 – in particular the fact that the three largest EU countries, Germany, France and the UK, have reported the most reliable figures – suggest that there is some link between the size of countries and the reliability of their data, with the largest countries reporting better data. It is not clear whether this is because of institutional issues, economies of scale or critical mass, or because of some implication of the laws of large numbers (compilation errors have a visible impact in the accounts of small countries, while they may offset each other for large countries). In any case, the example of Italy – which appears much lower in the rankings than much smaller countries, such as Ireland and Finland – shows that the link is tenuous. The causes of the differences in reliability among Member States is a topic for future research.

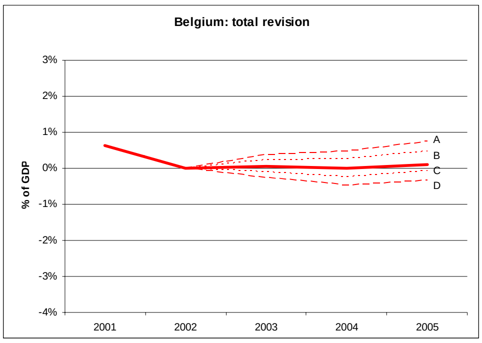
5 Inferring Future Revisions

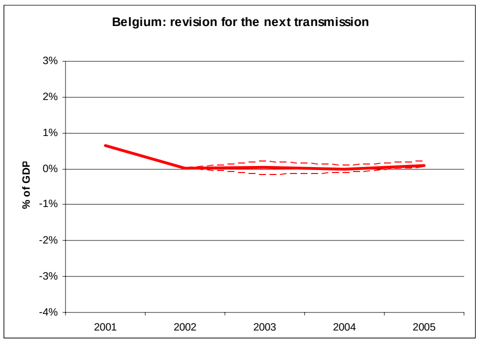
Information on the reliability of the deficit can be summarised in fan charts (see Figure 1). The fan charts are an efficient way of illustrating the uncertainty surrounding the deficit figures for the latest years. The charts depict how the deficit figures for recent years – which are not yet final – are susceptible to revision, based on the pattern of data revisions in the past. For every country, there are two charts. For each chart, the thick line depicts the latest available deficit ratios. Charts on the left-hand side show intervals of confidence for data to be reported in future data transmissions until data become final. That is, how data for 2003 will be reported by autumn 2007, data for 2004 will be reported by autumn 2008, and data for 2005 will be reported by autumn 2009. Data for years 2002 and earlier are no longer covered by the fiscal reporting and are, therefore, final. As expected, the confidence intervals around the thick line are wider for years which are still at an earlier stage of revision.

The intervals defined by the dotted and dashed lines correspond to confidence levels of 60% (a particularly low confidence level, uncommon in statistical analysis, which illustrates that deficit data are susceptible of large revisions) and 90%. This means that there is a probability of 60% that final deficit data will be between lines B and C, and 90% between lines A and D (see the chart for Belgium). In the context of fiscal surveillance in the EU, deficit-increasing revisions are of much more concern than deficit-reducing revisions; the unilateral confidence levels (i.e. probabilities that the government balance after the final revision will remain above lines C and D) are 80% and 95%, respectively.[23]

Charts on the right-hand side show intervals at a confidence level of 90% for the next data transmission. That is how data are susceptible of being revised in spring 2007.[24] The unilateral confidence level is 95%.

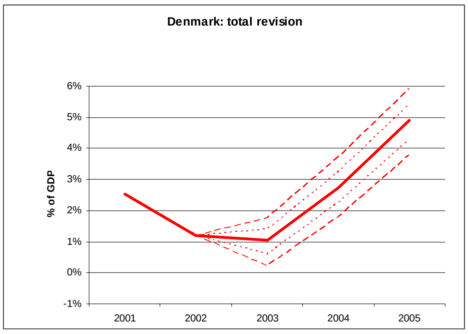
Figure 1: Fan charts illustrating the reliability of deficit statistics

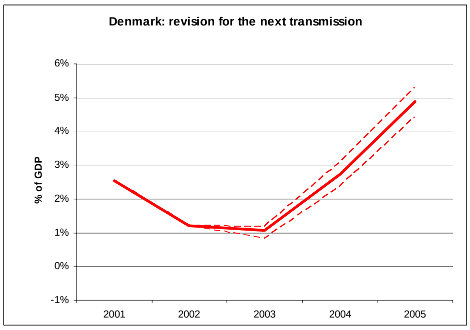


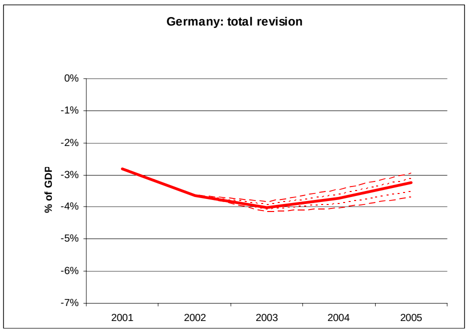


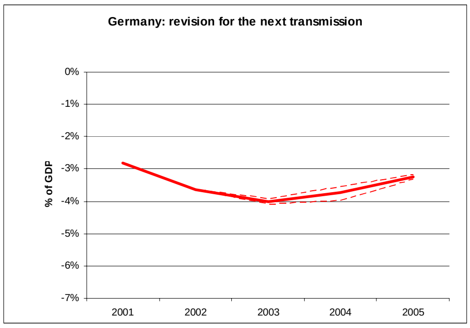
23. The intervals of confidence have been calculated on the basis of the average revision, the standard deviations in revisions from each data transmission to final, assuming that the revisions are normally distributed.

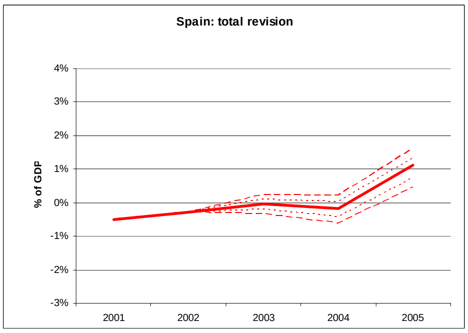
24. In spring 2007, there will be a major structural revision in the Danish and Swedish government accounts relating to the transitory period on the sectoral classification of pension schemes. Such a structural revision is not considered in the charts of Figure 1; it will reduce the Danish and Swedish surpluses by around 1% of GDP. See, for further details, European Commission (2006b, in particular Box 1.5).

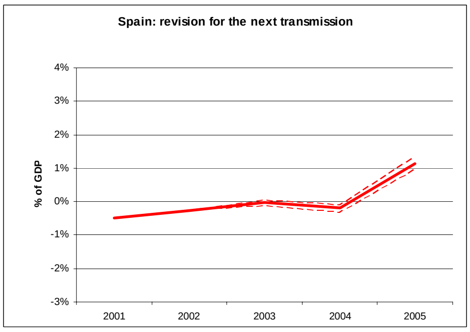

Denmark: total revision
% of GDP
6%
5%
4%
3%
2%
1%
0%
-1%
2001
2002
2003
2004
2005

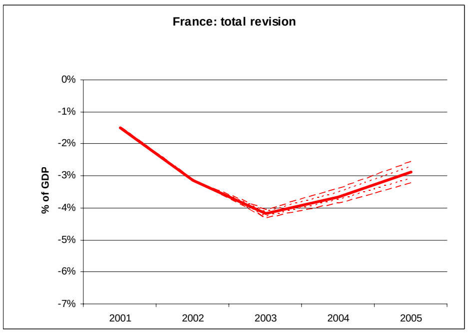

Denmark: revision for the next transmission
% of GDP
6%
5%
4%
3%
2%
1%
0%
-1%
2001
2002
2003
2004
2005

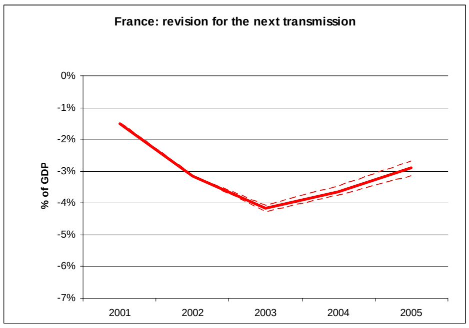

Germany: total revision
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005

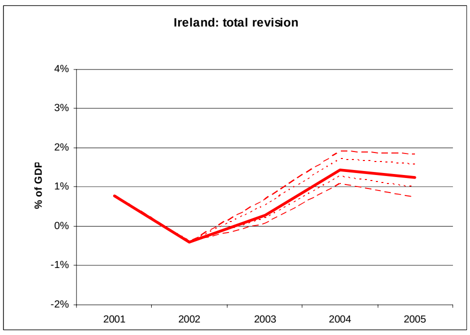

Germany: revision for the next transmission
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005

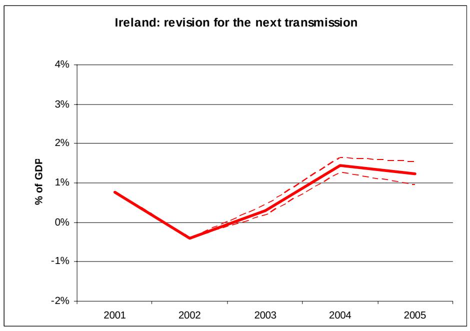

Spain: total revision
% of GDP
4%
3%
2%
1%
0%
-1%
-2%
-3%
2001
2002
2003
2004
2005

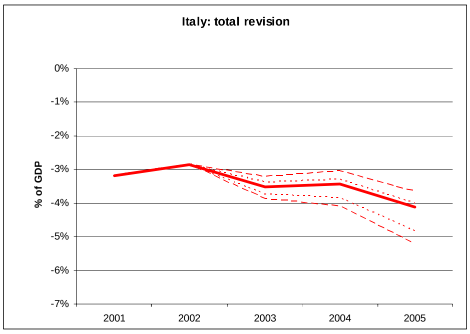

Spain: revision for the next transmission
% of GDP
4%
3%
2%
1%
0%
-1%
-2%
-3%
2001
2002
2003
2004
2005

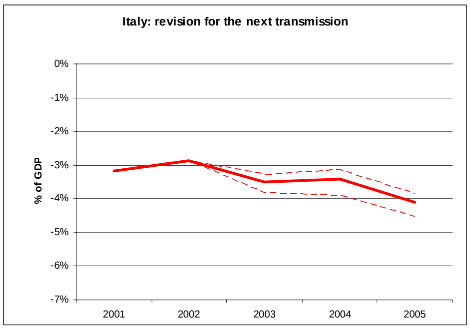

France: total revision
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005

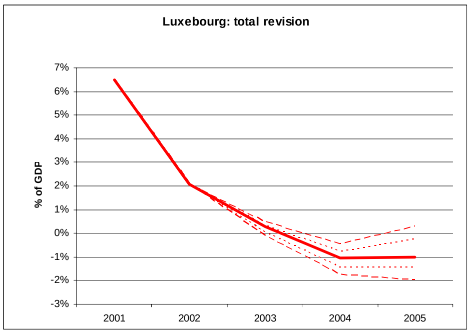

France: revision for the next transmission
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005

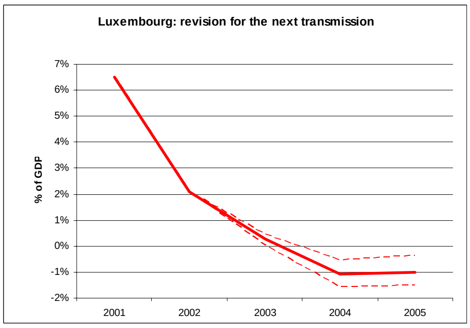

Ireland: total revision
% of GDP
4%
3%
2%
1%
0%
-1%
-2%
2001
2002
2003
2004
2005

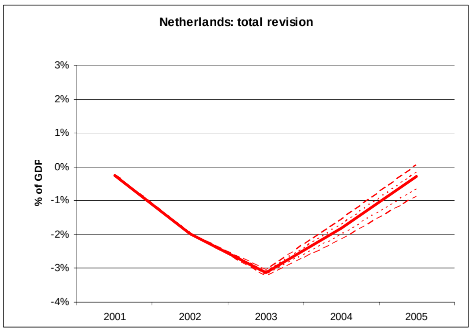

Ireland: revision for the next transmission
% of GDP
4%
3%
2%
1%
0%
-1%
-2%
2001
2002
2003
2004
2005

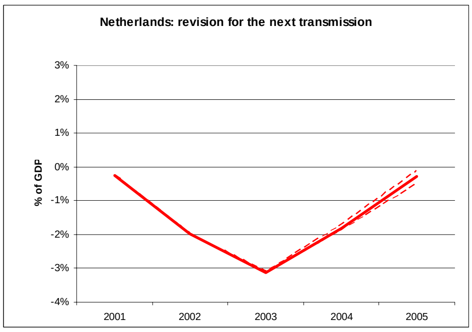

Italy: total revision
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005

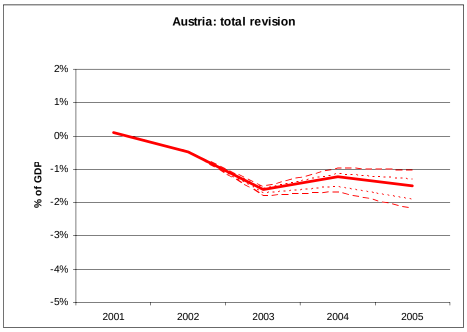

Italy: revision for the next transmission
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005

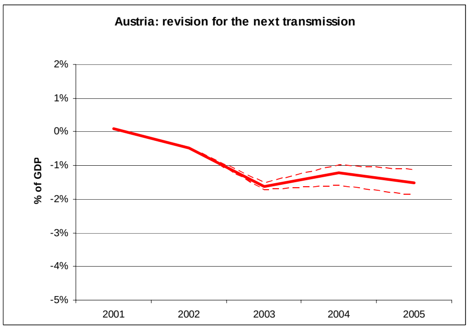

Luxebourg: total revision
% of GDP
7%
6%
5%
4%
3%
2%
1%
0%
-1%
-2%
-3%
2001
2002
2003
2004
2005

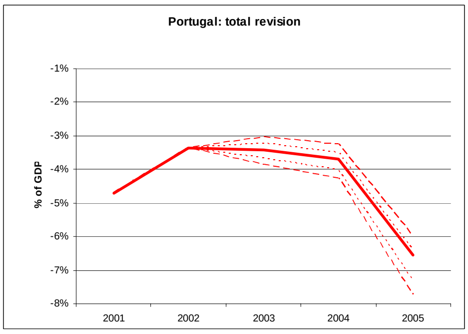

Luxembourg: revision for the next transmission
% of GDP
7%
6%
5%
4%
3%
2%
1%
0%
-1%
-2%
2001
2002
2003
2004
2005

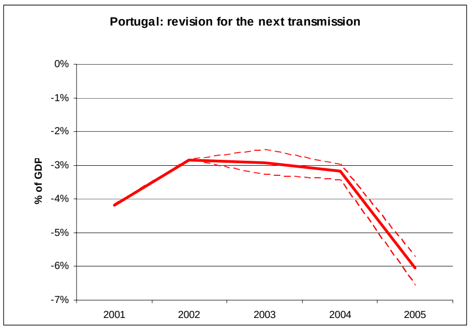

Netherlands: total revision
% of GDP
3%
2%
1%
0%
-1%
-2%
-3%
-4%
2001
2002
2003
2004
2005

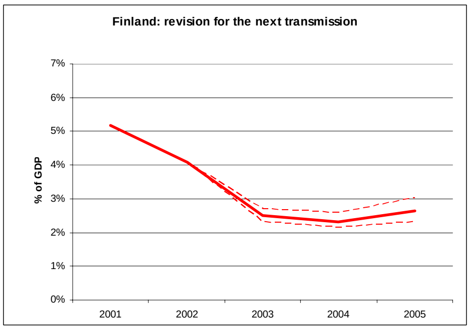

Netherlands: revision for the next transmission
% of GDP
3%
2%
1%
0%
-1%
-2%
-3%
-4%
2001
2002
2003
2004
2005

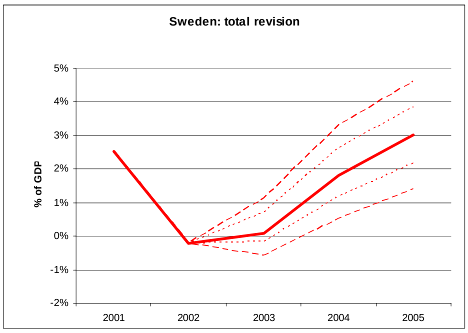

Austria: total revision
% of GDP
2%
1%
0%
-1%
-2%
-3%
-4%
-5%
2001
2002
2003
2004
2005

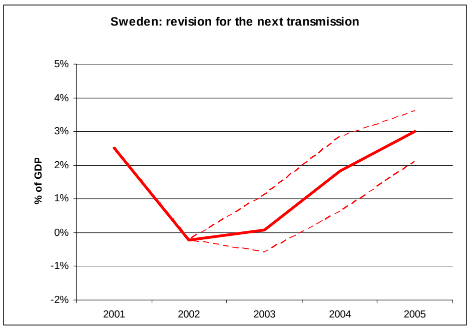

Austria: revision for the next transmission
% of GDP
2%
1%
0%
-1%
-2%
-3%
-4%
-5%
2001
2002
2003
2004
2005

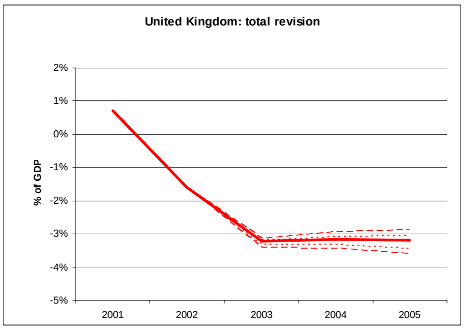

Portugal: total revision
% of GDP
-1%
-2%
-3%
-4%
-5%
-6%
-7%
-8%
2001
2002
2003
2004
2005

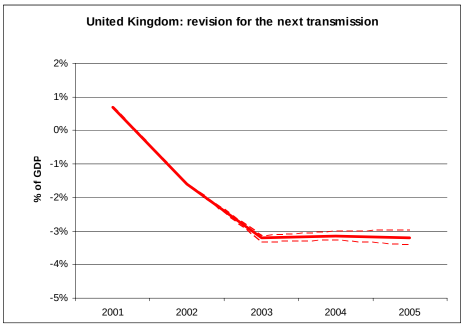

Portugal: revision for the next transmission
% of GDP
0%
-1%
-2%
-3%
-4%
-5%
-6%
-7%
2001
2002
2003
2004
2005











Note: This series of charts show the latest available deficit series (the thick line) and confidence intervals surrounding those figures. The dotted and the dashed lines show how the deficit figures are susceptible of revision until they become final (left-hand side charts). There is a probability of 60% that data will be comprised between lines B and C, and of 90% between lines A and D (see the first chart for Belgium). In terms of unilateral confidence levels, there is a probability of 80% that the deficit/surplus will not go below line C and of 95% that it will not go below line D. For the next data transmission (right-had side charts), the bilateral confidence level is 90%; the unilateral confidence level is 95%.

6 Comparing deficit and debt revisions

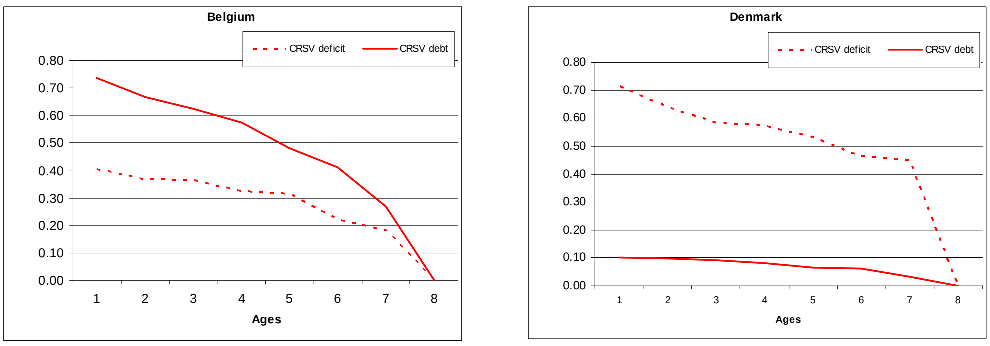
Debt figures are often assumed to be more reliable than deficit figures.[25] This is because accounting rules for debt are much simpler than those for deficits (debt is a cash-based concept which is directly observable and for which data are easier to compile than deficits which are calculated on an accrual basis), and because the deficit is a balancing flow that is influenced by all changes in other items of government accounts, while debt is a stock, which by nature is more inert than a flow. The idea that the deficit figures are less reliable also comes from the fact that the deficit has been a more politically sensitive variable.[26] Therefore, the governments have more interest in underestimating the deficit figures in the first data transmission, even if the data will be revised at a later stage. In this section, we test whether this view is well founded.

A direct comparison between the reliability indicators of the deficit and debt figures as those shown in sections 4 and 5 would be meaningless, because of differences in the size and nature of both concepts.[27] A meaningful comparison of the reliability of deficit and debt must take into account these differences. We suggest comparing the reliability of deficit and debt figures through two complementary approaches.

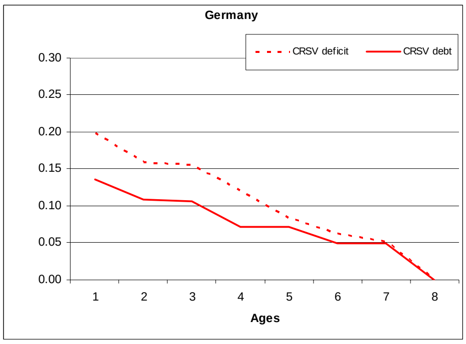
The first approach is to calculate a coefficient of reliability standardised by volatility (CRSV). We have defined the CRSV as the standard deviation of the revisions for every data age divided by the volatility of the underlying series. The volatility of each series (deficit and debt) is measured by the standard deviation of the yearly changes. The CRSV is a noise-to-signal ratio. The rationale of the CRSV is that volatile series are more difficult to compile in a timely way and therefore require larger revisions. A higher value in CRSV means less relative reliability.

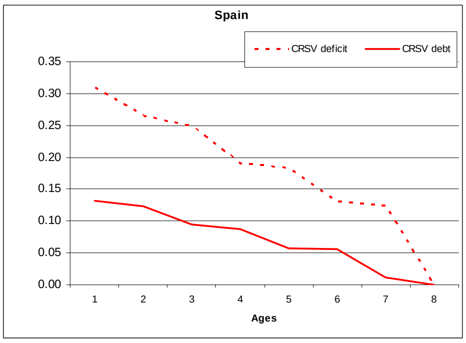
Charts in Figure 2 show the evolution of CRSV by age and country for deficit and debt figures. Since data of age 8 are defined as final, the CRSV for age 8 for all series is zero, by construction. Although the CRSV of deficit and debt can be compared for each age, the most useful comparison is for age 1, that is, for the first data transmissions. Debt figures have been more relatively reliable than deficit figures for Denmark, Germany, Spain, Ireland, Italy, Netherlands, Finland, Sweden and the United Kingdom. However, the reverse is true for Belgium, France, Luxembourg, Austria and Portugal.

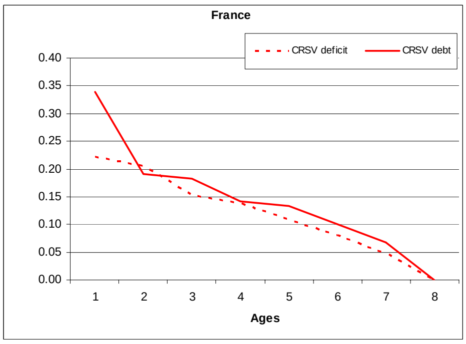
Figure 2: Coefficients of reliability of deficit and debt standardised by volatility



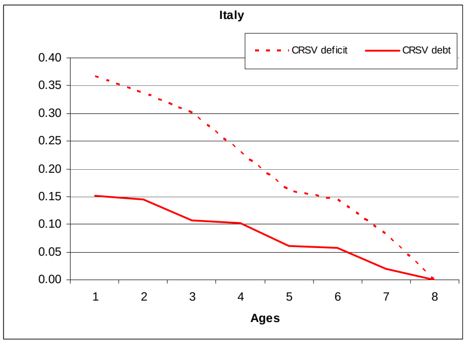
25. Although the government deficit/surplus is the main driver of the government debt, the yearly change in the government debt does not exactly correspond to the government deficit. The difference is known as stock-flow adjustment (or SFA). For further details on the components of the stock-flow adjustment, and a theoretical model on the authorities' behaviour in relation to deficits, debt and SFA in the presence of fiscal rules, see Buti et al (2006).
26. Balassone et al. (2003, 2006a and b), Buti et al. (2006) and Koen and van den Noord (2006).
27. Annex III provides a series of indicators on the revisions of debt data, similar to those shown above for the deficit.

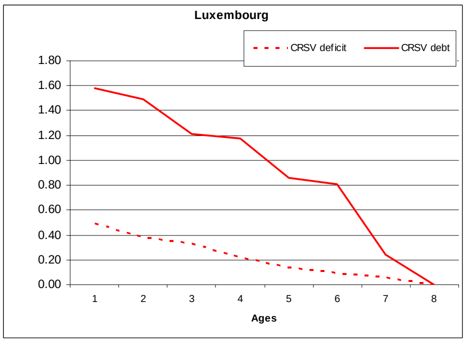

Germany
CRSV deficit
CRSV debt
Ages

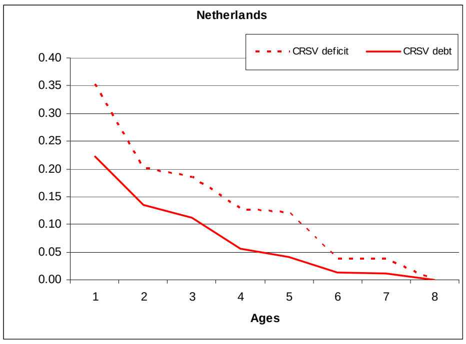

Spain
CRSV deficit
CRSV debt
Ages

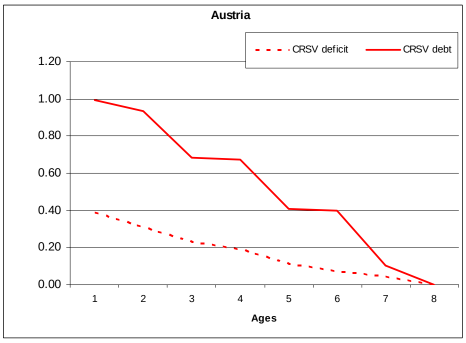

France
CRSV deficit
CRSV debt
Ages

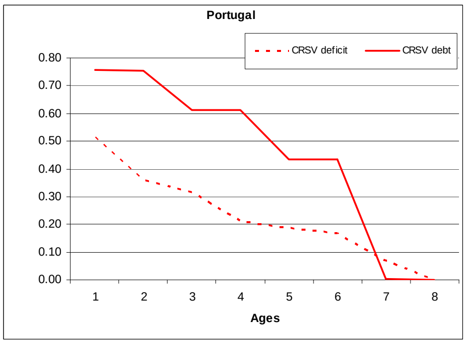

Ireland
CRSV deficit
CRSV debt
Ages

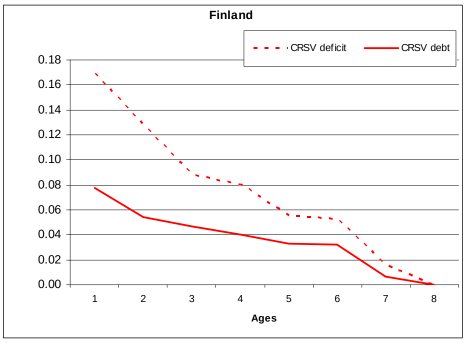

Italy
CRSV deficit
CRSV debt
Ages

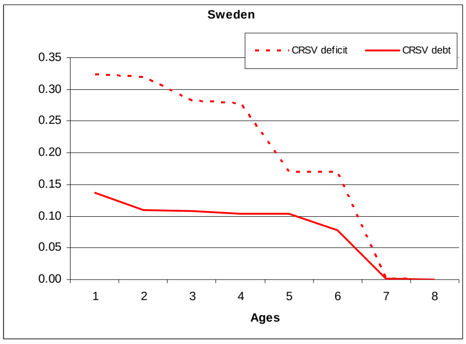

Luxembourg
CRSV deficit
CRSV debt
Ages

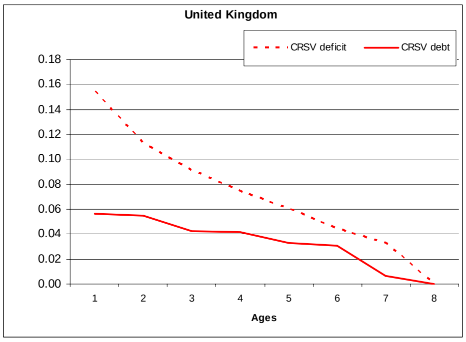

Netherlands
CRSV deficit
CRSV debt
Ages


Austria
CRSV deficit
CRSV debt
Ages


Portugal
CRSV deficit
CRSV debt
Ages


Finland
CRSV deficit
CRSV debt
Ages





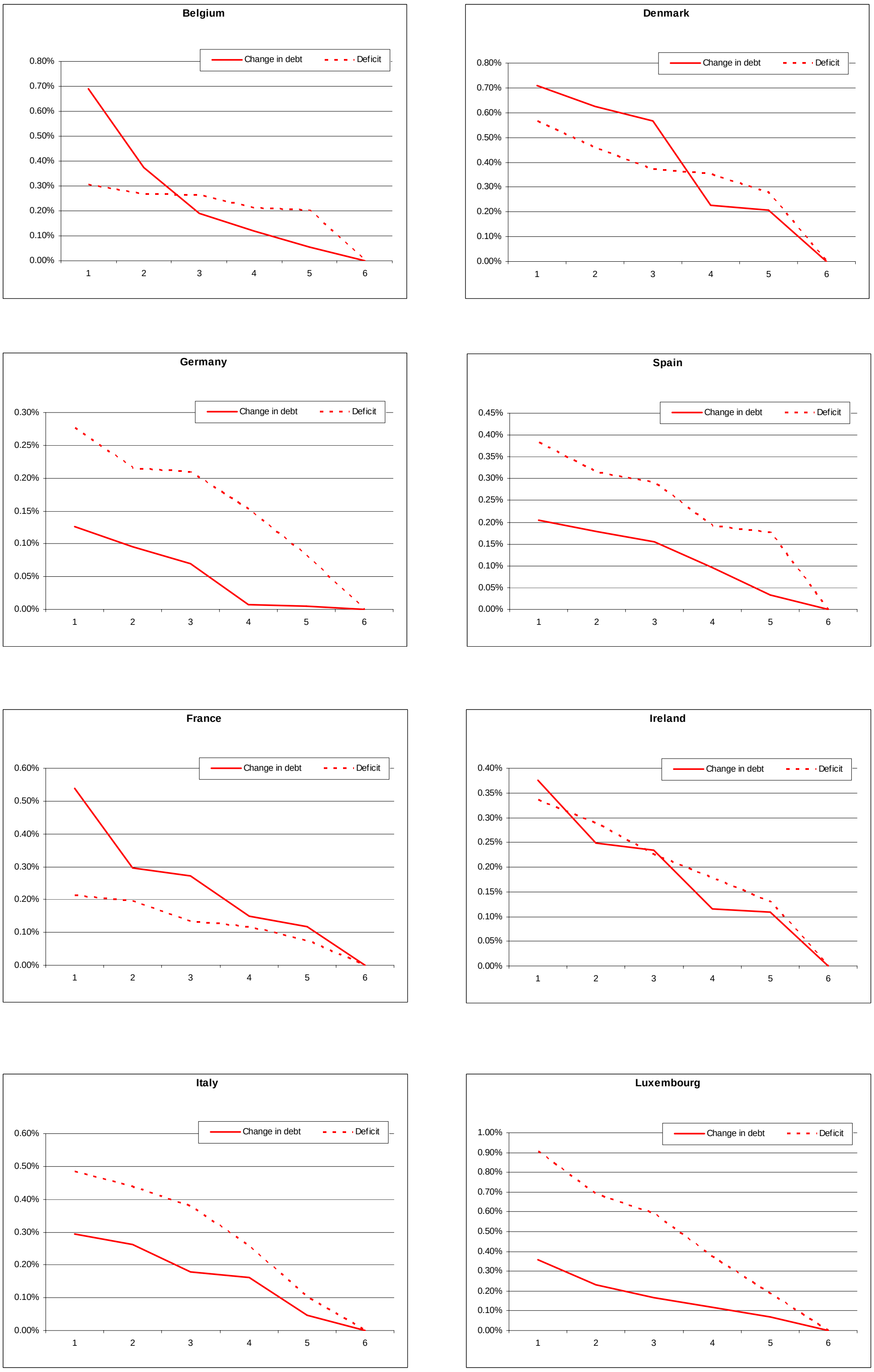
Note: This series of charts compare the coefficients of reliability standardised by volatility (CRSV) of deficit and debt ratios of different ages. The CRSV is calculated as the standard deviations of revisions divided by the standard deviation of year-on-year changes of the underlying series. A smaller coefficient for deficit (debt) indicates that deficit (debt) figures are relatively more reliable.

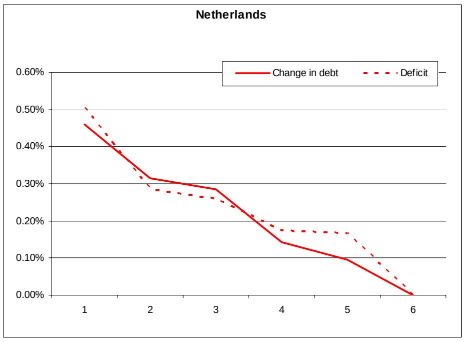
A second approach to compare the reliability of deficit and debt figures is to compare the revisions in deficit figures and the revisions in the yearly changes in debt, where the comparison is between two flows of similar sizes. This approach is the more relevant if one is interested in the debt developments to countercheck the plausibility of deficit figures. The charts in Figure 3 show the evolution of the standard deviations in the revisions of deficit and yearly change in debt figures. Since assessing changes in debt implies a comparison of two years' data reported at the same time, the analysis only concerns the first six data transmissions.

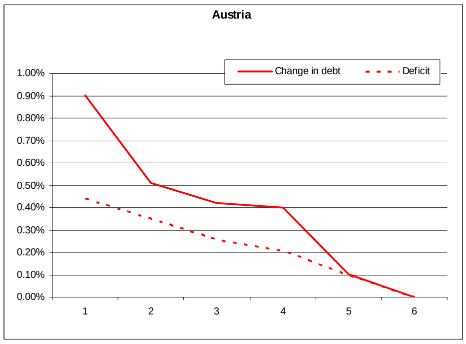
These charts have been complemented by F tests to establish whether the differences between the standard deviations of the deficit and change-in-debt revisions are statistically significant. The conclusions derived from the charts and the F tests are that, as far as data of age 1 are concerned, the figures on the yearly changes in debt are effectively more reliable than the deficit for Germany, Spain, Luxembourg, Italy and Portugal. Therefore, for these countries, the yearly change in debt may be used to countercheck the plausibility of deficit figures, and even to anticipate future deficit revisions. However, the results for Belgium, France, Austria and Finland go against the received wisdom, with their deficit figures more reliable than the debt figures. For, Denmark, Ireland, Netherlands and Sweden, the differences in the standard deviations are not statistically significant.

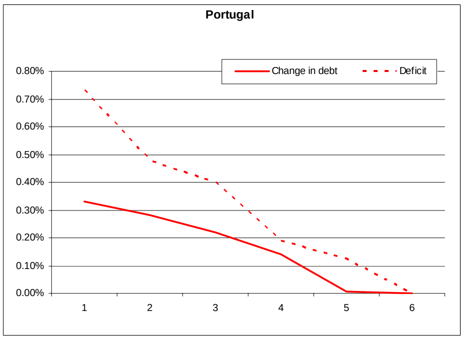
We conjecture that differences in the relative reliability of deficit and debt data come from the institutional arrangements for the compilation of deficit and debt figures, which in several countries are not compiled by the same institution, debt decentralisation and the size of government liabilities held by other government units (that is, liabilities which are consolidated when calculating the overall government debt).

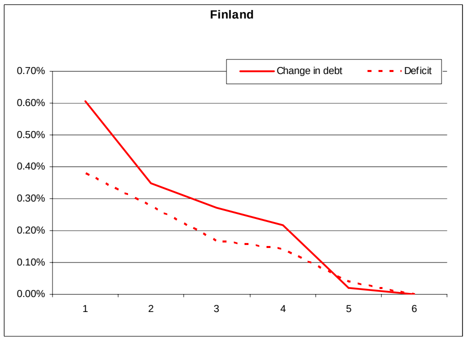
Figure 3: Comparing the reliability of deficit and change in debt data: evolution of the standard deviations of revisions

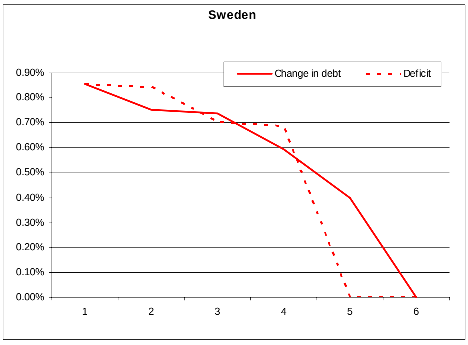


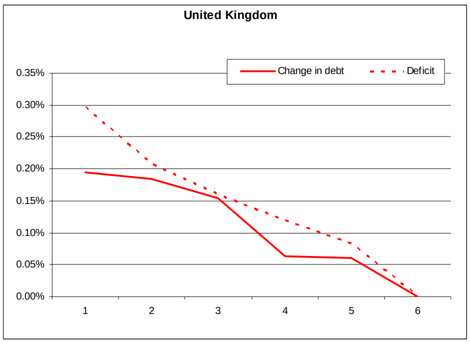












Note: This series of charts compares the evolution of the standard deviations of deficits and yearly changes in debt. When the yearly change in debt is more reliable (i.e. smaller standard deviation) than the deficit, debt figures can be used to assess the plausibility of, and anticipate revisions in, deficit figures.

7 Evolution of reliability over time

The process of compiling government accounts has changed quite considerably over the last 15 years. Technological advances in the automated collection and processing of data, the development of quarterly accounts and an increase in the awareness of the relevance of reliable government accounts may have led to an increase in the reliability of data. The increasing relevance of government accounts has also led to an upgrade in the quantity and quality of resources dedicated to compiling government accounts in Member States. Therefore, first outturns may be closer to final data than they used to be. At the same time, the shift from ESA79 to ESA95 may have led to an increase in the complexity of accounting rules – for example a wider application of the accruals principle – which would imply larger revisions. Moreover, the EU governments have participated more and more in complex transactions – such as securitisation or private-public partnerships – that raise accounting difficulties and often entail large revisions. The fact that Eurostat now scrutinises the accounts reported by Member States in much more detail than it used to do in the past may also have led to changes in the way data are compiled and revised.

In order to reach conclusions on the evolution of reliability over time for the deficit data, we have pooled data of the fourteen countries, and split the sample into two periods. The first period includes data corresponding to data transmissions under ESA79, that is from spring 1994 until autumn 1999; the second period includes data compiled under ESA95 since spring 2000.

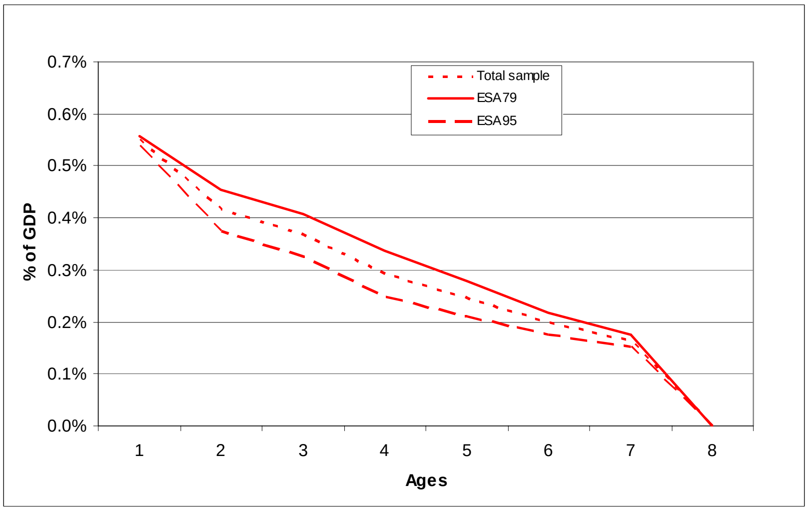
Figure 4 shows the evolution in the standard deviation in the revisions for the pooled sample and for the sub-samples ESA79 and ESA95. The standard deviations of the cumulated revisions from the first transmission to final data are practically the same under ESA79 and ESA95 while the data transmitted from the second to the sixth transmission appear more reliable under ESA95 than under ESA79. F tests confirm this finding. This indicates that figures compiled since 2000 under ESA95 converge quicker to final than data compiled prior to 2000. This faster convergence to final data may be connected with the fact that until 2000, several Member States used to revise the complete government expenditure and revenue account only once a year. Therefore, while one cannot conclude that ESA95 data are more reliable that ESA79 data, the conjecture that the shift from ESA79 to ESA95 has harmed the reliability of data is not confirmed by the evidence.[28, 29]

---

28. To check the robustness of our finding, we performed the same exercise for the group of countries that have ranked lower in terms of reliability (Netherlands, Italy, Denmark, Portugal, Luxembourg and Sweden; see the ranking in table 2). For this group, the standard deviation in revisions from age 1 to final data for ESA95 is slightly higher than for ESA79, while the reverse is true for other ages (2 to 7). However, the differences in the standard deviations are not significant for any age.

29. We have also compared the evolution of the average of revisions in absolute value under ESA79 and ESA95. The conclusions are the same as reported above for the standard deviation. Our finding that there has not been a general deterioration in the reliability in data reported prior and after the shift to ESA95 is not necessarily valid for each specific country. Given the small samples available, it is not possible to check whether the general finding is applicable to each individual country.

Figure 4: Evolution of reliability (standard deviation of revisions) of deficit under ESA79 and ESA95



Note: This chart compares the evolution of the standard deviations of the revisions in deficits between each given age and final data, when compiled under ESA79 and under ESA95. The conjecture according to which the shift from ESA79 to ESA95 has harmed the reliability of data is not confirmed by the evidence. Actually, data compiled since 2000 under ESA95 converge faster to final data.

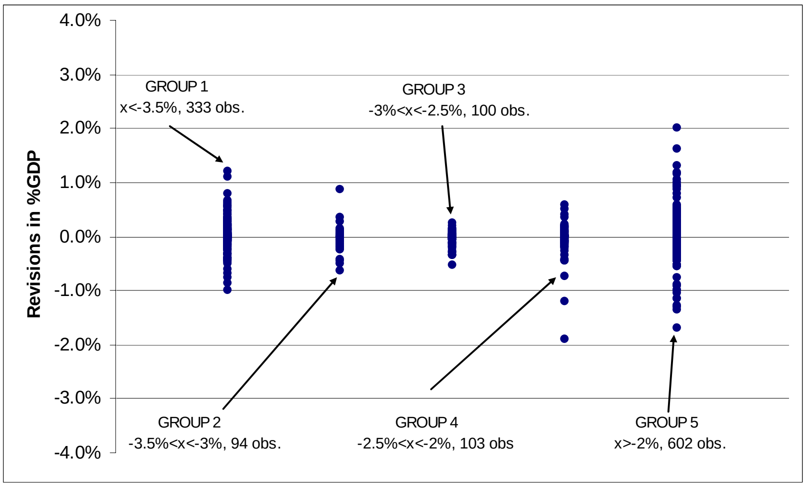
8 Comparing size of revisions and size of deficit

We have also investigated whether the size of the deficit has an impact on the revisions. From the data compiler viewpoint, there is no obvious reason for the size of the deficit to have an impact on the pattern of revisions. That is, there is no reason to expect that the errors that require correction and underlie future revisions are significantly different when the deficit is high or low, when one is in deficit or when one is in surplus. However, in the context of SGP and the related procedures, the reference value of 3% of GDP for the deficit plays a crucial role. Statistical authorities are certainly not insensitive to the political consequences of the revisions.[30] A revision which increases or reduces a surplus or that increases or reduces a large deficit does not have the same consequences as a revision that leads the deficit ratio to cross the bar of 3% of GDP.

To check whether the size of the deficit ratio has an impact on the pattern of revisions, we have pooled all revisions (i.e. of all countries and of all vintages) in a total of 1232 observations. This sample has been divided into five groups. The first group includes revisions affecting deficits which prior to the revisions are above 3.5% of GDP, while the second group covers revisions of deficits between 3.0% and 3.5% of GDP. The third group encloses the most sensitive revisions in the sense that they affect deficits between 2.5% and 3% of GDP; this group includes the cases where relatively small deficit-increasing revisions could imply crossing the 3%-of-GDP threshold. The fourth group includes revisions affecting deficits between 2.5% and 2.0% of GDP and the last one, revisions of deficits lower than 2.0% or surpluses. The sample distributed within the five groups is shown in Figure 5.

The figure suggests that the revisions of critical deficit figures – that is of deficit figures that are slightly below 3% of GDP – are smaller than the revisions when the government balance prior to the revisions is far from 3%. F tests confirm that the standard deviation of revisions in group 3 is significantly smaller than in any other group. Note, however, that revisions of group 3 are symmetric.

Figure 5: Relation between size of revision and size of deficit (-) or surplus (+)

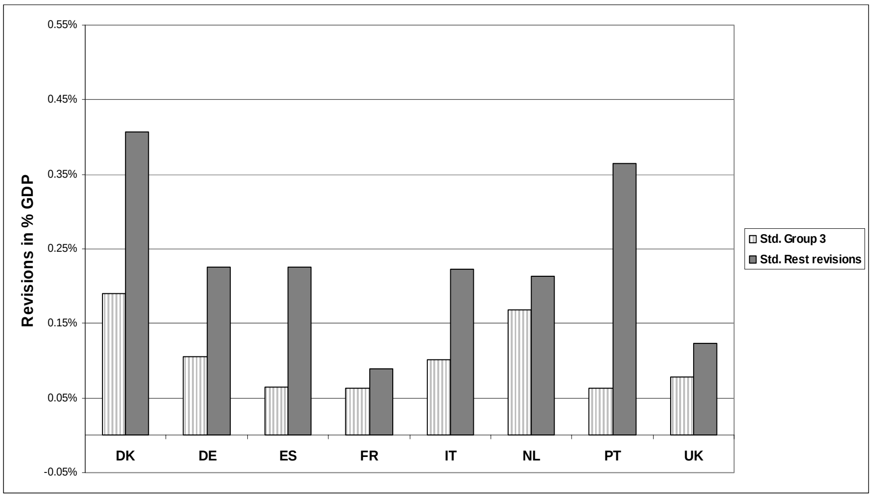


Note: The revisions in deficit figures have been plotted according to their size and the deficit (or surplus) prior to the revisions. The chart suggests that deficits slightly below the reference value of 3% of GDP (group 3) have been revised by less than when the deficit is above, or much below 3% of GDP.

30. For a synopsis of the recent debate on the institutional independence of data compilers, see European Commission (2006a, in particular section II.4.4.4).

Figure 6 also provides information on the standard deviation of revisions for group 3 (deficits prior to the revision slightly below 3% of GDP) and the other groups for a number of countries. The figure does not cover all countries, as several of them never reported deficits between 2.5 and 3% of GDP. It confirms the findings of Figure 5 that the revisions of deficits slightly below 3% of GDP are smaller than revisions in other circumstances.

Figure 6: Standard deviations of deficit revisions by country and size of deficits



Note: The chart shows the standard deviations of revisions for group 3 (deficits prior to the revision slightly below 3% of GDP) and the other groups for a number of countries. It confirms that revisions of deficit when the deficit prior to the revision is slightly below 3% (group 3) are usually smaller than for other deficit values. The chart does not cover all countries, as several of them never reported deficits between 2.5% and 3% of GDP.

There are several possible interpretations for Figures 5 and 6. A first interpretation is that deficit data are more carefully compiled when the deficit is close to 3% of GDP and, as a result, subsequent revisions are smaller. A more disturbing interpretation would be that statisticians refrain from revising their data with the same rigour when the deficit is very close to the reference value of 3% of GDP, as if they were afraid of the political consequences of revisions resulting in the reference value being crossed. If this second interpretation were correct, then one might expect larger revisions around the 3%-of-GDP mark very late – when data are no longer officially transmitted to the Commission for SGP purposes and therefore not covered by our analysis – or that data reach final status before all potential revisions have been implemented. However, further research would be necessary before drawing definitive conclusions on this topic.

9 Conclusions

The time frames for SGP-related policy decisions on the one hand, and of publication and revision of deficit and debt figures on the other, are quite different, the former being much faster than the latter. The first outturns of government accounts are subject to wider margins of error than the implementation of the fiscal surveillance procedure would require. On a number of occasions, this has generated stress and problems of credibility, yet clearly the fiscal surveillance procedures cannot simply be stopped in order to wait until final data are available.

The fact that it takes so long to get final data and that policy decisions are taken on the basis of preliminary information is not new and is not specific to the Stability and Growth Pact. In their decision-making process, policy-makers must keep in mind that the available information is not perfect and subject to revision. However, given that fiscal surveillance in the EU is primarily based on the evolution of specific indicators and the respect of specific ceilings or targets, it is necessary to take those errors explicitly into account. This means that those that decide on the SGP-related procedure must be aware that fiscal statistics are susceptible to relatively large revisions until very late and that the size of these revisions differs from country to country. Although the Commission cannot predict how reported data will be revised, it would be useful for it to make explicit what the uncertainty margins surrounding the headline figures are each time important steps in the SGP are taken. In some cases, it may make sense to postpone the abrogation of a procedure when the deficit ratio falls to below 3% of GDP, but remains so close to it that there is a significant risk that the deficit will have to be revised to above the reference value at a later stage, in particular if some other indicators suggest that an upward revision is likely. This proposal is consistent with the reformed Stability and Growth Pact, which puts the emphasis on durable consolidations.

This paper has looked into deficit and debt figures compiled and reported by 14 Member States (i.e. excluding Greece and the new Member States) since 1994 for the period 1990 to 2005. On the basis of the successive revisions in reported data, it assesses the reliability of budgetary statistics, derives a number of reliability indicators and infers the size of future data revisions. According to the standard deviations of revisions between the first outturn and final data, France, Germany and United Kingdom have reported the most reliable deficit/surplus figures. Denmark, Portugal, Luxembourg and Sweden are the countries that have had the largest dispersion of revisions. If one considers the average revisions, the two countries with largest deficit-increasing revisions have been Portugal and Italy.

The paper has also shown that for several countries, notably Germany, Spain, Luxembourg, Italy and Portugal, data on the yearly changes in debt have been more reliable than those on the deficit. For these countries, the yearly change in debt may be useful for counterchecking the plausibility of deficit figures and anticipating revisions in deficit figures. However, for several other countries, deficit figures have been more reliable than debt, or the difference is not significant.

We have not identified any significant difference in the reliability of the first outturns of deficit data when compiled under the former ESA79 system and ESA95. However, ESA95 converge more quickly to final than data compiled prior to 2000 under ESA79. A conjecture according to which the shift from ESA79 to ESA95 would have led to a general deterioration in the reliability of data is not confirmed by the evidence.

Our work has shown that the revisions of critical deficit figures – that is of deficit figures that are slightly below 3% of GDP – are smaller than the revisions when the government balance prior to the revisions is far from 3%. However, a more detailed analysis would be necessary before drawing conclusions on the causes of this result.

Data revisions are inevitable; in a sense, they are even desirable. All countries should put particular emphasis on the improvement of their statistical system, to improve the reliability and other quality dimensions of their data. In particular, those countries that have registered the largest data revisions should carefully analyse their revisions on a regular basis and consider whether changes in the statistical processes are needed so that the reliability of the first release improves.

Although we studied more than 1200 revisions, the available samples per country are relatively small. As a result, our results should be considered with prudence. However, that should not preclude an awareness of the distinct variations in reliability in the data compiled and reported by different countries.

References

ARUOBA, B. (2005). *Data Revisions are not Well-Behaved,* CEPR Discussion Paper No 5271, CEPR/EABCN No 21.

BALASSONE, F., D. FRANCO and S. ZOTTERI (2003). 'Fiscal Rules and Statistics'*,* paper presented at the *IIPF Congress,* Prague [also published as chapter 10 in Buti and Franco (2005)].

— (2006a). 'EMU Fiscal Indicators: A Misleading Compass?', *Empirica*, 33 (2), pp. 63-87.

— (2006b). *The Reliability of EMU Fiscal Indicators: Risks and Safeguards,* mimeo.

BIER, W. and H. AHNERT (2001). *De afruil tussen tijdigheid en nauwkeurigheid,* Economisch Statistische Berichten, No. 4299 (English version: *Trade-off between Timeliness and Accuracy*, available at: http://dsbb.imf.org/vgn/images/pdfs/bierahnert.pdf).

BRUNILA, A., M. BUTI and D. FRANCO (eds.) (2001). *The Stability and Growth Pact – The Architecture of Fiscal Policy in EMU*, Palgrave.

BUTI, M., and D. FRANCO (2005). *Fiscal Policy in EMU - Theory, Evidence and Institutions*, Edward Elgar.

BUTI, M. (2006). 'The use of government accounts for the implementation of the SGP, and the SNA review', presentation at the *Joint meeting of the Advisory Expert Group on National Accounts and National Accounts Users*, held at the ECB, Frankfurt am Main, 9 February 2006 (available at http://www.ecb.int/events/pdf/conferences/aeg/Buti2.pdf).

BUTI, M., J. NOGUEIRA MARTINS and A. TURRINI (2006). *From Deficits to Debt and Back: Political Incentives under Numerical Fiscal Rules,* CEPR Discussion Paper No 5809, (CESifo Economic Studies, forthcoming).

CARSON, C., and L. LALIBERTÉ (2002). *Assessing accuracy and reliability: a note based on approaches used in national accounts and balance of payment statistics,* IMF Working Paper 02/24.

DAMIA, V., and C. PICÓN AGUILAR (2006). *Quantitative Quality Indicators for Statistics – An Application to Euro Area Balance of Payment Statistics,* ECB Occasional Paper No 54.

DE VRIES, W. (2002). 'Dimensions of Statistical Quality - A discussion note about the quality initiatives of some international organisations', paper presented at *Inter-agency Meeting on Coordination of Statistical Activities*, New York, 17-19 September 2002, ref. SA/2002/6/Add.1.

DI FONZO T. (2005). 'Revisions in Quarterly GDP of OECD countries', paper presented at the *OECD Working Party on National Accounts meeting,* 11-14 October 2005*.*

EUROPEAN COMMISSION (2003). 'Fiscal Policies in EMU: 2003', *European Economy,* No 3.

— (2005). 'Fiscal Policies in EMU: 2005', *European Economy*, No 3.

— (2006a). 'Fiscal Policies in EMU: 2006', *European Economy*, No 3.

— (2006b). 'Convergence Report December 2006 - Technical Annex', (SEC (2006) 1570), *European Economy*, No 1.

EUROSTAT (2003). 'How to make a Quality Report', paper presented at the Working Group *"Assessment of quality in statistics"*, Luxembourg, 2-3 October 2003.

GIOVANNINI, E. (2004). 'The Quality of International Statistics: Outstanding Issues and Proposals for Improvement', in *Statistics and Their Use for Monetary and Economic Policy-Making*, ECB, Frankfurt am Main.

IMF (2003). *Data Quality Assessment Framework for Government Finance Statistics,* Statistics Department, July.

KOEN, V., and P. VAN DEN NOORD (2006). *Fiscal Gimmickry in Europe: One-off Measures and Creative Accounting,* in Wierts et al. (eds.) (2006).

KOPITS, G., and S. SYMANSKY (1998). *Fiscal Policy Rules,* IMF Occasional Paper, 162.

LALIBERTÉ, L., and D. DEFAYS (2006). 'Comparison of the IMF's Data Quality Assessment Framework and the European Statistical System Quality Approach: An Update', paper presented at the *European Conference on Quality in Statistical Surveys*, Cardiff, 24-26 April 2006.

LALIBERTÉ, L., W. GRÜNEWALD and L. PROBST (2004). *Data Quality: A Comparison of IMFs Data Quality Assessment Framework (DQAF) and Eurostat's Quality Definition,* mimeo.

OECD (2003). *Quality Framework and Guidelines for OECD Statistical Activities*, STD/QFS(2003)1.

SAEBOE, H. V., J. BYFUGLIEN and R. JOHANNESSEN (2003). 'Quality Issues at Statistics Norway', *Journal of Official Statistics,* 19:3 pp. 287-303.

TANZI, V. (2006). *Fiscal Policy: When Theory Collides with Reality,* CEPS Working Document No 246 (June).

WIERTS, P., S. DEROOSE, E FLORES and A. TURRINI (eds.) (2006). *Fiscal Policy Surveillance in Europe*, Palgrave Macmillan.

I- Data management

I.1 Available sample

The available sample consists of data reported to the Commission by Member States in the context of the excessive deficit procedure. Each data transmission (also called a vintage) covers the previous four years. The data transmissions also include forecasts for the current year, which are not taken into consideration in this paper. The sample contains data for 16 years: 1990 to 2005, reported between March 1994 and October 2006. Until the end of 2005, data used to be reported by 1 March and 1 September each year. From the beginning of 2006, the reporting dates moved to 1 April and 1 October.[31] Therefore, deficit and debt figures of year t are reported for the first time in spring t+1, and then successively revised every 6 months until autumn t+4.

Table I.1 shows the available sample for the United Kingdom; other Member States' data have a similar shape. Note that the number of available ages is not the same for each year, since each year has not been subject to the same number of data transmissions. Data for years 1993 to 2002 were reported eight times, that is, from the first outcome until final data; for 2003 six times: the seventh transmission will take place in spring 2007 and the eighth in the autumn 2007; for 2004 there have been four transmissions to date, while data for 2005 have so far been transmitted only twice. Data for 1990 were subject to two transmissions only (they dropped out of the data transmission after being reporting twice, in spring and autumn 1994), data for 1991 were transmitted four times, and data for 1992 six times.

Table I.1: Example of basic data sample – deficit data reported by the United Kingdom

Deficit (-) figures	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Mar-94	-8300	-15900	-37800	-48300												
Sep-94	-8165	-15127	-36601	-48664												
Mar-95		-15127	-36601	-49365	-46063											
Sep-95		-14995	-37575	-49093	-45586											
Mar-96			-37575	-49093	-45382	-41742										
Sep-96			-37540	-49310	-45380	-40358										
Mar-97				-49310	-45380	-39052	-32311									
Sep-97				-49660	-45532	-38854	-35971									
Mar-98					-45613	-38592	-34805	-14377								
Sep-98					-45613	-38592	-34805	-16142								
Mar-99						-40032	-32970	-14823	5302							
Sep-99						-40032	-32970	-15292	3940							
Mar-00		ESA 95 starts					-32909	-16294	2549	10728						
Sep-00							-32909	-16294	3810	11835						
Mar-01								-16294	3810	11420	41453					
Sep-01								-16924	3810	11420	40439					
Mar-02									3458	10333	39086	8803				
Sep-02									1992	10322	37477	7622				
Mar-03										10322	37477	7881	-14058			
Sep-03										9092	36136	7154	-15401			
Mar-04											36106	6987	-16277	-34945		
Sep-04											36351	6716	-18074	-36049		
Mar-05												6716	-18074	-36977	-36893	
Sep-05												7234	-16658	-36640	-36139	
Mar-06													-16817	-36715	-38660	-44727
Sep-06													-17039	-36257	-37641	-39732

As the case of the UK in Table I.1 illustrates, each country sample contains 104 observations. However, there are in fact 8 different sub-samples for each country corresponding to the 8 ages of data with 13 observations for every age.

31. In the paper, we disregard this change in the transmission deadlines and will simply refer to the spring and autumn data transmissions.

I.2 Levels and ratios

Revisions in the deficit and debt ratios can occur because of revisions in the numerators (deficit/debt) and denominators (GDP). In this paper, however, we study only the revision in the numerators. We have disregarded the fact that the nominal GDP figures are also subject to revision and that such revisions have an impact on the deficit/debt ratios. This is for two reasons, both pragmatic. First, routine revisions in the GDP figures rarely have a significant impact on the deficit ratios. This can be illustrated with a simple numerical example. Assume a deficit of 3% of GDP; if the nominal GDP for a given year is revised upwards by 1% – which, for a routine revision, would be a very large, and extremely rare, revision – the deficit ratio would fall by 0.03% points only. Of course, there are also larger revisions in GDP of a structural nature. For example in autumn 2005, the GDP levels of the Netherlands, Spain and Portugal were revised upwards by around 4%.[32] However, by their very nature, structural revisions are not relevant to this analysis, as they do not help us infer the size of future revisions.

A second reason for not considering revisions in the denominator is connected with the differences in the deficit and GDP compilation. The deficit is mainly compiled by an additive procedure. This means that the identification and correction of a routine error in the deficit for a given year t does not necessarily have any implications for the deficit of year t+1. The compilation of GDP is mainly a multiplicative process; the identification and correction of a routine error in the GDP level of a given year t usually leads to similar revisions for the subsequent years. As a result, the reliability of GDP levels is much more complex to measure, and analyses of GDP revisions often focus on the growth rates.[33]

Although we are only interested in the revision in the numerators, there is a need to standardise the deficit/debt levels and the respective revisions across time and countries. Therefore, all calculations are as a percentage of GDP; to simplify, we use the GDP series available in the European Commission's AMECO databank.

I.3 Synthetic indicators of reliability

For each data transmission, there are data of different ages and with different revision histories. In spring t, data for t-1 are transmitted for the first time, data for t-2 are transmitted for the third time, data for t-3 are transmitted for the fifth time and data for t-4 are transmitted for the seventh time. Each data transmission thus contains data of different ages: ages 1, 3, 5 and 7 in spring, and 2, 4, 6 and 8 in autumn. We aim at calculating reliability indicators for each age, that is, to estimate how these figures are susceptible of revision before reaching final status in age 8. Data of age 2 are expected to be more reliable than date of age 1, data of age 3 are expected to be more reliable than data of age 2, etc. Data of age 8 are final.

For years that have been subject to eight transmissions, there are seven revisions between two consecutive vintages. These revisions can be treated as random variables. Therefore, there are seven variables for the deficit and debt figures of each country for which a small number (10 to 13) of observations are available. This is obviously a very small sample for inference purposes; a more appropriate sample which would allow better inference would need to contain around 30 revisions for each age of data. However, a sample of such size will only be available beyond 2020.

---

32. We should also note a revision in the Greek GDP series by 26% as announced in September 2006. At the time of writing, the new series has not yet been validated by Eurostat (see Eurostat News Release N° 139/2006 of 23 October 2006). In any case, Greece is not systematically considered in this paper.

33. On revisions in GDP figures see, Aruoba (2005) and Di Fonzo (2005).

Our aim is to infer the behaviour of these variables representing the revisions in order to calculate synthetic indicators of reliability. These variables, corresponding to the revisions of the different ages, can be cumulated to calculate total revisions from any given age to final data. It is possible to apply the central limit theorem and the law of large numbers (Lyapunov theorem) assuming that the general conditions of this theorem are complied and that the variables are normally distributed. Given the small size of the available sample by country, we were not able to perform the usual normality tests, such as the Lilliefors test. The Shapiro-Wilks test, which is appropriate for samples smaller than 50 observations, has been applied and it shows that for 28 variables of a total of 196 (98 for deficit and 98 for debt) the null hypothesis of normal distribution is not rejected. The Lilliefors test shows that the complete sample (pooling data of the 14 countries) is normally distributed, a result which is a consequence of the law of large numbers.

In order to set a procedure to calculate variances, and also as a previous condition of the Lyapunov theorem, it is crucial to test the independence of the variables. If variables are normally distributed and are independent, the characteristic function of the sum is the characteristic function of a variable normally distributed with a mean equal to the sum of the means and variance equal to the sum of variances. The Pearson product moment correlation coefficient shows that 57% of the total variables on a country-by-country basis are independent; the Spearman rank correlation coefficient allows 86% of variables to be considered independent. Therefore, we have accepted the independence of the series.

Furthermore, we have carried out equality tests for mean and variance on a basis of an ANOVA model with a single factor. The test shows that it is not possible to reject the equality of means of the revisions over the previous vintages, but allows us to clearly reject the null hypothesis for equality of variance across variables.

To calculate the average revision and the standard deviation of these revisions for each age, one could simply compare the series of data for each age with final data. However, because of the small size of the samples, the heterogeneity of samples (for several years, we do not have data of all ages) and the structural breaks that need to be removed from the sample – namely the break because of the shift from ESA79 to ESA95 – another technique must be used. If we simply compared data of age 1 with data of age 8, we could only work with data for 1993 to 1995 and 1999 to 2002. The data transmissions for 1990 to 1992 would be lost as the data of earlier ages are not available; data for 1996 to 1998 would also be lost because there was a structural break (shift to ESA79 to ESA95) in the transmission of spring 2000; and finally data for 2003 to 2005 would also be lost as the final ages are still missing.

We have circumventing the heterogeneity of the sample by calculating the averages and variances of revisions in data between each consecutive ages of data (data of age 1 are compared to data of age 2; data of age 2 are compared with data of age 3, and successively). The average revision from each age until age 8 is calculated by adding the successive average revisions together. Assuming that the variables are independent, as discussed above, that is, assuming that the revisions linked to an age are independent of revisions of other ages, the standard deviations are calculated by adding up the unbiased estimators of variances of each variable, that is, of each age revision, and then calculating the square root.

II- Reliability indicators and structural revisions

II.1. Denmark

Denmark used to classify 2nd-pillar pension schemes in government. In September 2004, pursuant to the Eurostat decision of 2 March 2004, Denmark reclassified these pension schemes in the financial sector; this led to a fall in the government surplus of around 1% of GDP and an increase in government debt. In March 2005, Denmark returned to the previous sectoral classification in line with a transitory period granted by Eurostat on 23 September 2004. Denmark will once again have to change the sectoral classification of pensions schemes in spring 2007.[34]

When calculating the reliability indicators for Denmark, the deficit and debt revisions of autumn 2004 and spring 2005 for each of the concerned years were put together and the total revision (the two revisions largely offset each other) was evenly distributed among the relevant data ages.

II.2. Portugal and Sweden

As discussed in section 5, the inference on future revisions is meaningful only if the sample contains only routine revisions and structural revisions have been removed. This was not systematically attempted in this paper, given the small size of the available sample – there is the risk of throwing the baby out with the bathwater – and the difficulty of identifying structural revisions.

However, to illustrate the impact that structural revisions may have on data, Figure II.1 shows fan charts (confidence level of 90%) for Portugal and Sweden, two of the countries most affected by structural breaks. The dotted lines are as in the main text of the paper – that is, the only structural break identified and removed corresponds to the shift from ESA79 to ESA95. The long dashed lines define the fan chart after removing a number of country-specific structural breaks.

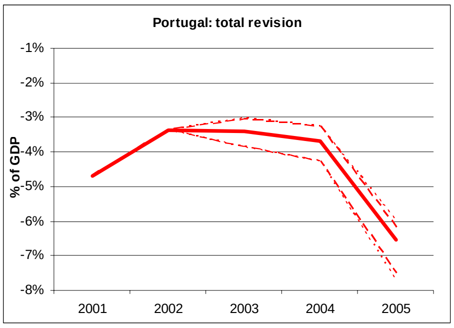
The structural breaks removed from the sample of Portugal are the 5th revision for 1999, the 3rd revision for 2000 and the 1st revision for 2001, which took place after the so-called Constâncio-I report of summer 2002. Those deficit revisions concerned mainly the recording of capital injections and taxes. For Sweden, the revisions removed are the 6th revision for 2000, the 4th revision for 2001 and the 2nd revision for 2002 which concerned changes in the recording of taxes.

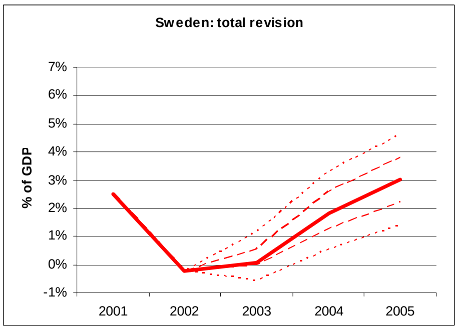
Although the number of revisions removed from the Portuguese and Swedish samples is the same, the impact on the reliability indicators as summarised in the fan charts is quite different. For Portugal, the impact is negligible; for Sweden, the fan becomes much narrower. These two cases illustrate that quantitative indicators on the reliability of data must be interpreted carefully, in particular for countries that have undergone structural changes in the compilation of their data.

---

34. See Eurostat News Releases N°30/2004 of 2 March 2004 and N° 117/2004 of 23 September 2004.

Figure II.1: Effect of structural breaks for Portugal and Sweden on the confidence intervals for the deficit/surplus





Note: The charts illustrate how the removal of structural revisions affects the reliability indicators for Portugal and Sweden. While for the former, the impact is negligible; the confidence interval becomes much narrower for Sweden. These two cases illustrate that prudence is needed when interpreting the reliability of data, in particular for countries that have undergone major changes in the compilation of their data.

III- Reliability indicators for debt

The analysis for deficit figures shown in sections 4 and 5 can be replicated to debt figures. Tables III.1 and III.2 show Member States ranked according to the standard deviations of debt data revisions and the average revision, while table III.3 shows the average of debt data revisions in absolute value for all countries.[35]

Information on the reliability of the debt can also be summarised in fan charts (Figure III.1), as we did for the deficit in section 5. The bands delimited by the dotted lines show the confidence intervals at a confidence level of 90% (or 95% for unilateral confidence levels).

Table III.1: Member States ranked by the standard deviation of revisions in debt (% of GDP)

	Revisions cumulated from						
	1st transmission to final	2nd transmission to final	3rd transmission to final	4th transmission to final	5th transmission to final	6th transmission to final	7th transmission to final
1	DE(0.33%)	DE(0.27%)	DE(0.26%)	NL(0.16%)	NL(0.12%)	NL(0.03%)	SE(0.01%)
2	ES(0.53%)	NL(0.38%)	NL(0.31%)	DE(0.17%)	DE(0.17%)	DE(0.12%)	PT(0.01%)
3	UK(0.54%)	FI(0.39%)	FI(0.33%)	FI(0.29%)	FI(0.23%)	FR(0.21%)	NL(0.03%)
4	FI(0.56%)	FR(0.41%)	ES(0.38%)	FR(0.30%)	ES(0.23%)	ES(0.22%)	ES(0.05%)
5	NL(0.62%)	ES(0.50%)	FR(0.39%)	ES(0.36%)	IT(0.28%)	FI(0.23%)	FI(0.05%)
6	IT(0.68%)	UK(0.53%)	UK(0.41%)	UK(0.40%)	FR(0.28%)	IT(0.26%)	UK(0.07%)
7	FR(0.72%)	SE(0.64%)	IT(0.48%)	IT(0.45%)	UK(0.31%)	UK(0.29%)	IT(0.09%)
8	SE(0.79%)	IT(0.65%)	SE(0.62%)	SE(0.60%)	LU(0.58%)	SE(0.45%)	DE(0.12%)
9	IE(1.03%)	IE(0.99%)	LU(0.81%)	LU(0.79%)	SE(0.60%)	LU(0.54%)	FR(0.14%)
10	LU(1.06%)	LU(1.00%)	IE(0.95%)	IE(0.88%)	IE(0.85%)	IE(0.64%)	LU(0.16%)
11	DK(1.82%)	DK(1.76%)	AT(1.54%)	DK(1.49%)	AT(0.92%)	AT(0.90%)	AT(0.23%)
12	BE(1.96%)	BE(1.78%)	DK(1.65%)	AT(1.52%)	DK(1.20%)	BE(1.10%)	DK(0.59%)
13	AT(2.24%)	AT(2.11%)	BE(1.66%)	BE(1.53%)	BE(1.28%)	DK(1.13%)	IE(0.60%)
14	PT(2.42%)	PT(2.41%)	PT(1.96%)	PT(1.95%)	PT(1.39%)	PT(1.39%)	BE(0.71%)

Note: The table shows the standard deviations of revisions in debt figures between any given data transmission and final data. The first column corresponds to the revisions between the first outturn for year t reported in spring t+1 and final data for the same year as transmitted in autumn t+4. A higher value in standard deviation of the revisions means more dispersion in deficit data for the same period over different data transmissions and a worse position of the country in terms of reliability.

35. Greece has not been systematically considered in the tables and charts. The standard deviation of revisions in debt data in Greece from the 1st transmission to final data has been 5.78% of GDP and the average revision is +4.17% of GDP.

Table III.2: Member States ranked by the average of revisions in debt (% of GDP)

	Revisions cumulated from						
	1st transmission to final	2nd transmission to final	3rd transmission to final	4th transmission to final	5th transmission to final	6th transmission to final	7th transmission to final
1	PT(-1.22%)	PT(-1.27%)	PT(-0.78%)	PT(-0.81%)	PT(-0.39%)	PT(-0.40%)	FR(-0.06%)
2	SE(-0.24%)	SE(-0.20%)	SE(-0.24%)	SE(-0.20%)	SE(-0.22%)	SE(-0.15%)	LU(-0.04%)
3	BE(-0.23%)	UK(-0.19%)	UK(-0.12%)	UK(-0.14%)	UK(-0.07%)	UK(-0.09%)	NL(-0.02%)
4	UK(-0.17%)	FI(-0.04%)	BE(0.01%)	BE(-0.08%)	NL(0.02%)	BE(-0.03%)	FI(-0.02%)
5	FI(-0.16%)	BE(-0.01%)	FI(0.03%)	NL(0.00%)	FR(0.02%)	NL(-0.01%)	UK(-0.01%)
6	NL(0.04%)	FR(0.07%)	FR(0.04%)	FI(0.04%)	FI(0.04%)	FR(0.02%)	SE(0.00%)
7	DE(0.15%)	NL(0.09%)	NL(0.07%)	DE(0.05%)	DE(0.05%)	DE(0.03%)	PT(0.01%)
8	FR(0.24%)	DE(0.13%)	DE(0.11%)	FR(0.08%)	LU(0.10%)	FI(0.06%)	IT(0.02%)
9	LU(0.52%)	LU(0.45%)	LU(0.30%)	LU(0.26%)	BE(0.14%)	ES(0.12%)	ES(0.02%)
10	ES(0.57%)	ES(0.47%)	ES(0.37%)	ES(0.28%)	ES(0.15%)	LU(0.13%)	DE(0.03%)
11	IE(0.59%)	IT(0.52%)	IT(0.37%)	IT(0.33%)	IT(0.17%)	IT(0.16%)	AT(0.03%)
12	DK(0.67%)	DK(0.54%)	DK(0.50%)	IE(0.46%)	AT(0.33%)	IE(0.24%)	IE(0.15%)
13	IT(0.70%)	IE(0.65%)	IE(0.54%)	DK(0.51%)	IE(0.36%)	DK(0.30%)	DK(0.18%)
14	AT(1.20%)	AT(1.19%)	AT(0.72%)	AT(0.71%)	DK(0.44%)	AT(0.33%)	BE(0.19%)

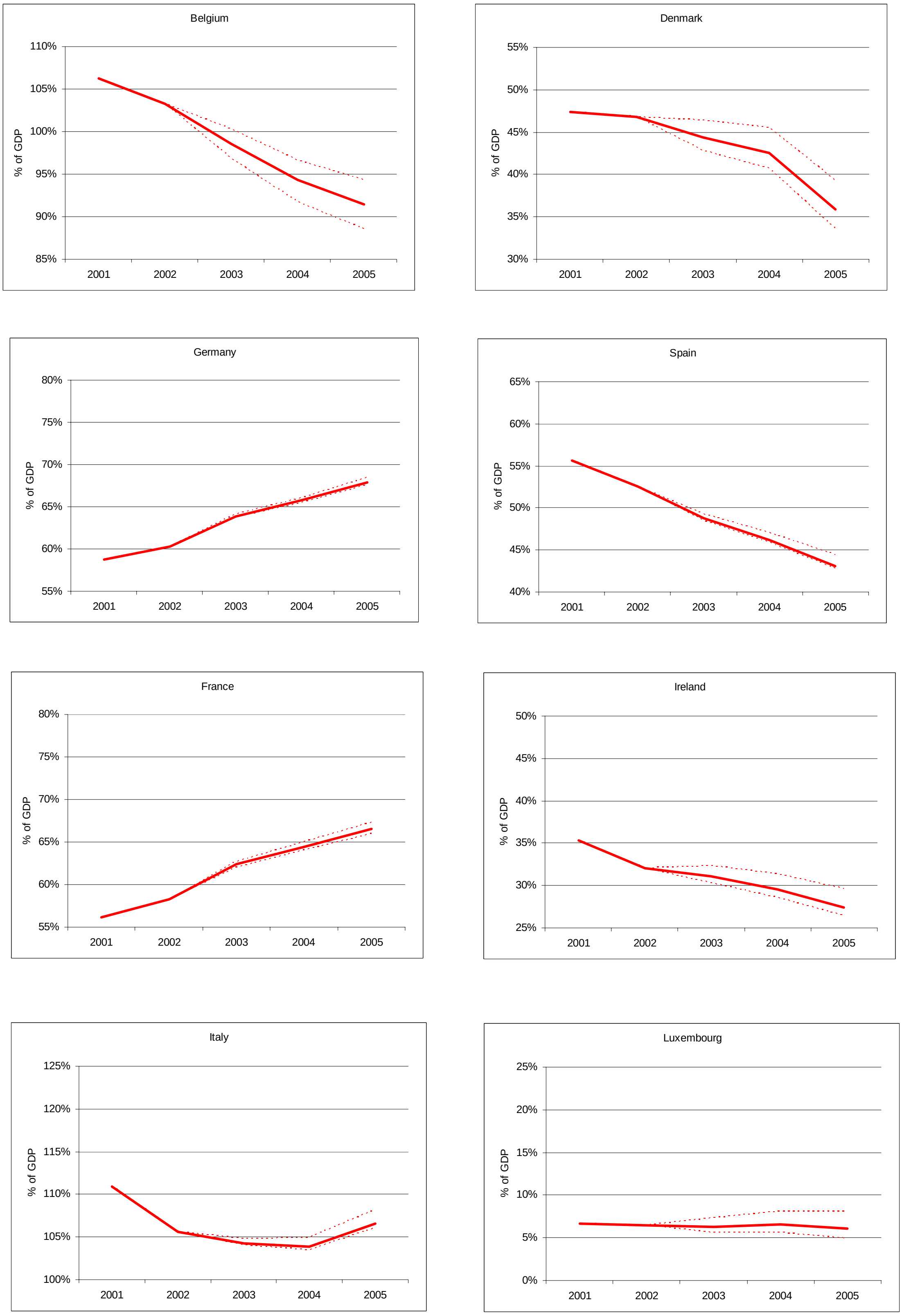
Note: The table shows the average revisions in debt figures between any given data transmission and final data. The first column (revisions from age 1 to age 8) corresponds to the revisions between the first outturn for year t reported in spring t+1 and final data for the same year as transmitted in autumn t+4. A negative (positive) value means an expected increase (reduction) in the debt.

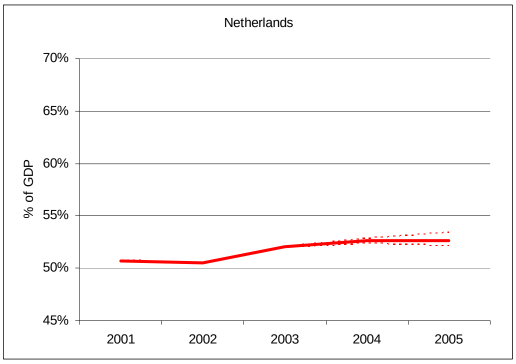
Table III.3: Member States ranked by the average of revisions in debt in absolute values (% of GDP)

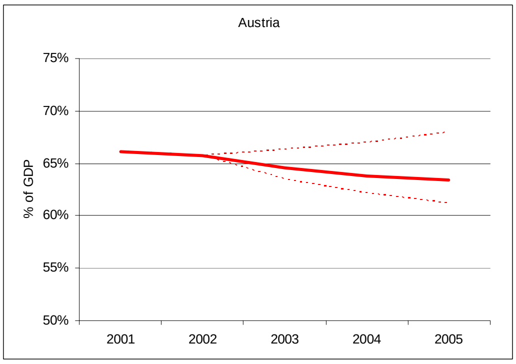
	Revisions cumulated from						
	1st transmission to final	2nd transmission to final	3rd transmission to final	4th transmission to final	5th transmission to final	6th transmission to final	7th transmission to final
1	DE(0.28%)	DE(0.17%)	DE(0.15%)	DE(0.08%)	DE(0.08%)	NL(0.02%)	SE(0.00%)
2	UK(0.54%)	SE(0.44%)	FI(0.29%)	NL(0.16%)	NL(0.10%)	DE(0.04%)	PT(0.01%)
3	SE(0.73%)	FI(0.44%)	UK(0.33%)	FI(0.18%)	FI(0.10%)	FI(0.09%)	FI(0.02%)
4	FI(0.74%)	UK(0.45%)	NL(0.35%)	UK(0.27%)	UK(0.18%)	UK(0.13%)	NL(0.02%)
5	ES(0.75%)	NL(0.50%)	SE(0.37%)	SE(0.31%)	ES(0.19%)	ES(0.15%)	ES(0.02%)
6	NL(0.86%)	FR(0.55%)	ES(0.45%)	FR(0.33%)	FR(0.26%)	SE(0.15%)	UK(0.04%)
7	FR(0.90%)	ES(0.65%)	FR(0.47%)	ES(0.36%)	SE(0.29%)	FR(0.16%)	DE(0.04%)
8	IT(1.05%)	IT(0.88%)	LU(0.61%)	LU(0.50%)	IT(0.30%)	LU(0.23%)	IT(0.05%)
9	LU(1.11%)	LU(0.90%)	IT(0.63%)	IT(0.54%)	LU(0.30%)	IT(0.23%)	LU(0.05%)
10	IE(1.49%)	IE(1.31%)	PT(1.00%)	AT(0.90%)	PT(0.44%)	AT(0.40%)	FR(0.08%)
11	PT(1.67%)	PT(1.53%)	AT(1.06%)	PT(0.91%)	AT(0.49%)	PT(0.44%)	AT(0.09%)
12	DK(1.93%)	DK(1.59%)	IE(1.15%)	IE(0.92%)	DK(0.70%)	IE(0.47%)	DK(0.18%)
13	AT(2.15%)	AT(1.63%)	DK(1.34%)	DK(1.06%)	IE(0.80%)	DK(0.56%)	BE(0.23%)
14	BE(2.41%)	BE(1.84%)	BE(1.49%)	BE(1.23%)	BE(0.82%)	BE(0.60%)	IE(0.36%)

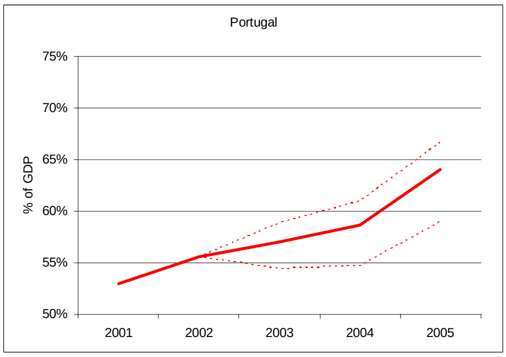
Note: The table shows the average revisions in debt in absolute values between any given data transmission and final data. The first column corresponds to the absolute revisions between the first outturn for year t reported in spring t+1 and final data for the same year as transmitted in autumn t+4. A higher value in this indicator means less reliable figures.

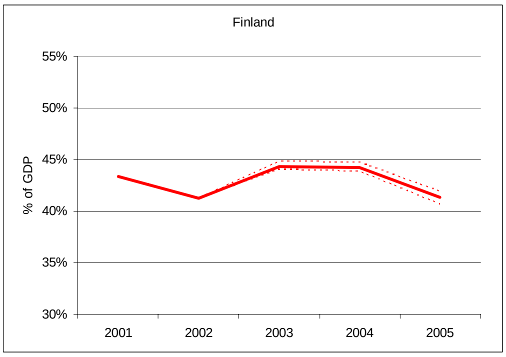
Figure III.1: Fan charts illustrating the reliability of government debt figures.

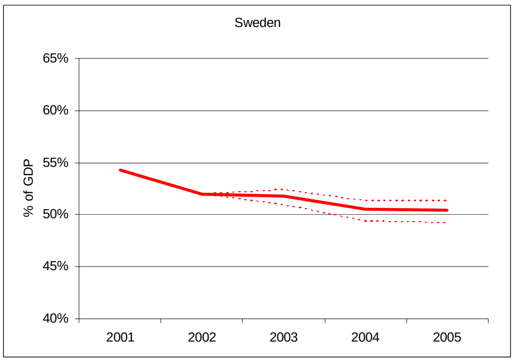


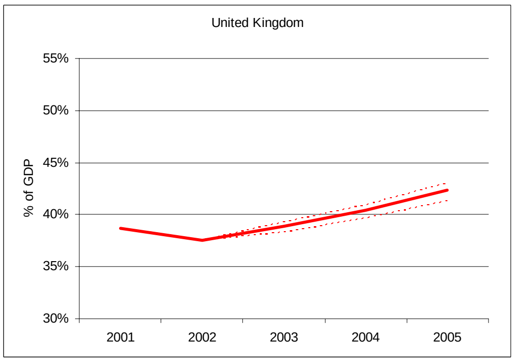












Note: This series of charts show the latest available debt series (the thick line) and confidence intervals surrounding those figures. The intervals defined by the dotted lines correspond to bilateral confidence levels of 90%. The unilateral confidence levels are 95%.

BANCO DE ESPAÑA PUBLICATIONS

STATISTICAL NOTES

1 DEPARTAMENTO DE ESTADÍSTICA Y CENTRAL DE BALANCES: Registering financial interemediation services on the nacional accounts as of 2005. (The Spanish original of this publication has the same number.)

2 DEPARTAMENTO DE ESTADÍSTICA Y CENTRAL DE BALANCES: Valuation of shares and other equity in the Financial Accounts of the Spanish Economy. (The Spanish original of this publication has the same number.)

3 DEPARTAMENTO DE ESTADÍSTICA Y CENTRAL DE BALANCES: Registering financial interemediation services on the nacional accounts as of 2005. Addendum. (The Spanish original of this publication has the same number.)

4 LUIS GORDO MORA AND JOÃO NOGUEIRA MARTINS: How reliable are the statistics for the Stability and Growth Pact?

BANCO DE ESPAÑA
Eurosistema

Unidad de Publicaciones
Alcalá, 522; 28027 Madrid
Telephone +34 91 338 6363. Fax +34 91 338 6488
e-mail: publicaciones@bde.es
www.bde.es