FIRSTRUN Deliverable 2.4

Revisions of the cyclically adjusted budget balance:
How large are they and how do they impact the European fiscal framework?

Abstract

The aim of this study is to analyse the revisions of the cyclically adjusted balances of the EU15 countries in order to learn whether the turn of the cycle is particularly prone to error and to evaluate whether the EU's fiscal framework is well equipped for dealing with potential revisions. The analysis is conducted with real-time fiscal data published by the European Commission for the years 2003-2016. Our study finds that the cyclically adjusted balance forecasts as well as in-year estimates are heavily revised in the following years. Moreover first ex-post data are still significantly revised in the following years. We find no strong evidence to support the notion that these revisions are larger at the turn of the cycle and thus conclude that revisions are of a systematic nature. The revisions are large enough to cause significant differences in the ex-ante and ex-post assessments by the European Commission on whether member states achieved their structural targets. This study elaborates on the built-in flexibility in the Stability and Growth Pact (SGP), on which the Commission can capitalise to avoid unwarranted ex-post sanctioning. Moreover, the SGP entails several safeguards against revisions negatively affecting the policy advice given to member states, particularly with regard to sanctioning.

Author:
Matthias Busse, CEPS

Table of Contents

List of Figures

List of Tables

List of abbreviations

CAB	Cyclically Adjusted Balance
CAPB	Cyclically Adjusted Primary Balance
CSR	Country-specific Recommendation
DBP	Draft Budgetary Plan
EA	Euro Area
EC	European Commission
ECB	European Central Bank
EDP	Excessive Deficit Procedure
EMU	Economic and Monetary Union
ESM	European Stability Mechanism
ESI	European Structural and Investment
FRED	Federal Reserve Economic Data
GDP	Gross Domestic Product
IMF	International Monetary Fund
MTO	Medium-term Objective
NBP	National Budget Plan
OECD	Organisation for Economic Co-operation and Development
OG	Output Gap
ROG	Representative Output Gap
SB	Structural Balance
SCP	Stability and Convergence Programme
SDP	Significant Deviation Procedure
SGP	Stability and Growth Pact
TSCG	Treaty on Stability, Coordination and Governance
WEO	World Economic Outlook

1. Introduction

The great recession, which followed the financial crisis of 2008, led to sharply rising debt-to-GDP ratios within the EU, prompting a set of reforms of the Stability and Growth Pact (SGP) aimed at encouraging member states to abide by the rules. In combination with the newly introduced Treaty on Stability, Coordination and Governance (TSCG), which came into force in January 2013, these should ensure that debt levels are brought down to a sustainable level. While policy-makers have emphasised the need for fiscal consolidation, they have advocated that it should be conducted in a counter-cyclical manner. Ultimately, the so-called structural balance took centre stage, which represents the general budget balance adjusted for the cycle,[1] as it expresses the fiscal stance corrected for the position of the business cycle.

The advantage of using structural variables within a set of fiscal rules is that the consolidation process is smoothed over the cycle and thus is less disruptive to economic activity. The disadvantage of employing structural variables is that they rely on the measurement of the output gap, which cannot be observed but has to be estimated. Hence the structural balance estimate may in real-time differ significantly from a later estimation, firstly because more data on the entire cycle become available, i.e. the real-time effect, and secondly since the output gap is not observable and may be re-estimated. If these revisions are large, it may lead to very different assessments ex ante compared to those performed ex post. Consequently revisions may prove that policy advice based on real-time data was inappropriate with regard to the cyclical position. Moreover, in extreme cases, the estimation error may trigger unwarranted sanctions or insufficient sanctions. Over the long run, this may jeopardise debt sustainability if the revisions show a persistently downward bias.

The extent to which this is the case depends, in the end, on how large the revisions are and in which direction they tend to move. In this paper we will show the magnitude of the revisions and their properties in order to assess if they constitute a serious problem for the SGP. Specifically we will establish whether the revisions are clustered at the turn of the cycle, which can have the effect of rendering revisions negligible in normal years, and causing the outliers at the peaks and troughs to drive the size of revisions. And if that is the case, the fiscal framework needs to be equipped with safeguards against the negative impact of revisions at turn of cycle. Conversely, if the revisions are systematic, we need to ascertain whether the SGP is capable of dealing with these revisions appropriately, i.e. if revisions are acknowledged, accounted for within the framework, which limits the probability of ill-fitting advice and sanctioning.

In quest of policy responses to these dilemmas, this paper is structured as follows: First we give an overview of the fiscal framework in order to establish when revision to the cyclically adjusted balance (CAB) could matter and which data are consulted in the assessment of compliance within the SGP. Secondly we will provide a literature review on real-time fiscal data analysis with an emphasis on the statistical properties. The third section describes the data employed and the methodological approach. Thereafter we provide an overview on the magnitudes of the revisions and highlight particularities. This analysis will be complemented by our findings whether revisions are structurally different at cyclical turns. With this knowledge in mind, we will assess the extent to which the revisions may lead to wrong policy advice and sanctioning – in particular, drawing on the first section and elaborating in which ways the SGP framework already mitigates the potential negative effects of revisions.

---

[1] The structural balance also deducts all 'one-off' items from the budget, in contrast to the cyclically adjusted balance.

2. Fiscal Framework

This section describes the European fiscal framework to show how structural fiscal targets have taken centre stage. The aim is to determine at what point in time the European Commission, in conjunction with the Council, assesses structural balances or improvements in the structural balance. Moreover, this section establishes which 'real-time' data are evaluated when the Commission asks for additional fiscal effort and sanctions are triggered.

2.1. Historical background

The inception of Economic and Monetary Union (EMU) provided new impetus for closer EU integration. One important milestone was the 1997 introduction of the Stability and Growth Pact. It is anchored in EU law, foremost in Articles 121 and 126 TFEU and Protocol 12 to the Treaty. It was designed to restrict excessive net lending of the member states and to install prudent fiscal policy in order to prevent the accumulation of debt to unsustainable levels. (European Commission, 2012a)

At its core the original SGP included two explicit targets: a debt level lower than 60% of GDP and an annual headline budget deficit of at most 3% of GDP. In case a country had already passed the 60% threshold, it should take actions to reduce the debt at an adequate pace. The adjustment path was not individually calculated for each country but was of a generic (one-size-fits-all) nature. While structural balances already featured in the SGP, the headline deficit target of 3% of GDP took the centre stage. (European Commission, 2012b)

In 2005 Germany and France were in violation of the 3% deficit target and pushed for a reform of the SGP and gained the needed support. Thereafter, the SGP emphasised the need to take the cyclical position into account. New exception clauses allowed member states to be temporarily exempted from the rules in the event they experienced an adverse economic shock. Moreover, the new SGP required an explicit (benchmarked) adjustment path towards the 60% debt level to be calculated on a country-by-country basis, i.e. taking into account the specificities of each economy. The country-specific adjustment path was expressed in the structural balance position – the CAB net of one-off budgetary items (non-recurrent revenues or expenditures, e.g. bank bailouts). (European Commission, 2016a)

In the wake of the financial crisis of 2008, debt levels had skyrocketed across the EU and a second reform phase to the SGP had been launched. The so-called Two-Pact and Six-Pact, which entered into force in 2011 and 2013, respectively, sought to strengthen the rules by introducing additional targets, reporting requirements and more prominence of sanctions (European Commission, 2016a).

2.2. The European Semester

For this paper it is of particular importance to verify when and which 'targets' are evaluated against which data and at what point the European Commission provides advice on national budgetary adjustments. In 2011 the fiscal surveillance framework took the form of the so-called European Semester (see Appendix 1).

Each April, member states submit their Stability and Convergence Programmes (SCPs), which detail the budget plans, underlying assumptions and risk, growth forecasts as well as estimates of the cyclically adjusted budget balance. These are to be provided for the preceding year (t-1) and until the three-years ahead (t+3). The forecasts are submitted with a complementary sensitivity analysis to highlight potential risks attached to the precision of such forecasts. (European Commission, 2012b)

In May of each year, the European Commission publishes its spring forecasts. On the basis of these data and those published in the SCPs the European Commission provides an assessment whether the budget plans laid out in the SCP are sufficient to achieve the fiscal targets. Thereafter the Commission produces its Country-Specific Recommendations (CSRs) detailing violations, risks and measures that would improve the fiscal stance.

In November the Commission issues an opinion on the Draft Budgetary Plan (DBP). In this opinion the Commission assesses whether the budget plan deviates from the (approved) budget plan in the SCP and, if the SCP had been previously deemed incompatible with the SGP rules, whether the new planned budget has addressed the shortcomings in the SCP budget plan. If the member state's budget is in violation of the SGP, the Commission can call for another revision. In December the member states adopt their final budgets. (European Commission, 2016a)

Hence the European institutions can influence the budget-making process throughout the entire year, but the most important instance is the evaluation of the SCP in spring when the first outturn data for the previous year are available and when a formal procedure can be launched, the first step towards sanctions. Which year(s) is examined to determine a violation depends on whether the country is not in compliance with the preventive or corrective arm of the SGP. (European Commission, 2016a)

2.3. The preventive arm of the SGP

The preventive arm of the SGP was introduced to keep government debt levels on a sustainable path over the economic cycle while allowing the automatic stabilisers to function as desired. Each member state is to achieve a so-called medium-term objective (MTO), which is expressed in terms of a structural balance stance. Country-specific MTOs are revised every three years but can even be revised more regularly in case of major reforms or changing economic circumstances (European Commission, 2016b). In case a member state's debt level is below 60%, it merely has to ensure that the level remains stable; however, if the country has already surpassed the threshold or faces future sustainability risks, then a structural budget should be chosen, which drives the debt level gradually down. In short, the MTO will be more ambitious the more precarious the debt sustainability is. There are three different MTOs which each fulfil a special requirement since all them are binding, effectively the strictest value is the one that matters (see European Commission, 2016a for more information). Overall an MTO is regarded as having been achieved if the deviation from it was less than 0.25% GDP.

In a first step the Commission establishes whether last year's budget (headline and cyclically adjusted) fulfils the requirements. In a second step the budget plan for the current and next year are scrutinised. Forecasts for t+2 and t+3 are only qualitatively assessed. If the country is found to have been at its MTO and predicted to continue to do so in the current and next year, no further steps are taken. If the MTO (or the adjustment path target) had not been met in the previous year (or as an average over the previous two years), a warning is issued. This can lead to a so-called significant deviation procedure (SDP), the first step towards sanctions. (European Commission, 2016a).

If a member state is on the adjustment path towards the MTO, it is asked, as a benchmark, to achieve a structural improvement of at least 0.5% GDP per annum. If the member state has accumulated a debt level beyond 60% of GDP, the required adjustment will be larger and if the economic environment (size of the output gap) is unfavourable, it will be smaller (see figure 1).

Figure 1: Required fiscal effort by position of the cycle
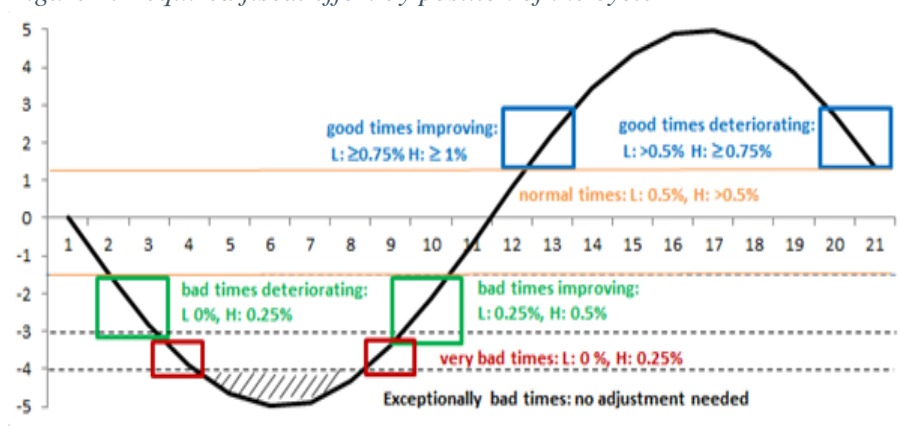
Source: Angela D'Elia, 2015.
Note: "L" stands for low debt and "H" for high debt.

If the structural improvement was not achieved, the Commission proposes new structural (improvement) targets and a deadline. Moreover the Commission may launched an SDP according to Article 121(4). A deviation of 0.5% (of GDP) from the adjustment path in the past year or a deviation of 0.25% (of GDP) on average over the two past years is deemed sufficient to trigger an SDP. For euro area (EA) member states, the SDP can result in a sanction requiring the country to deposit 0.2% of the previous year's GDP in an interest-bearing account. This sanction may be converted into a non-interest bearing deposit if the country also enters into an excessive deficit procedure (EDP) (see next section).

Besides the structural balance target, which is the focus on the paper, a second target has been introduced in the preventive arm: the expenditure growth benchmark. This indicator is calculated as the growth rate of government expenditures (excluding interest payments, cyclical unemployment spending, spending on EU programmes and smoothed for investment over a four-year time horizon) and compared to the potential growth rate of GDP[2]. If only one of the two indicators (structural balance or expenditure benchmark) is red-flagged, the grounds for this judgement is evaluated – one breach is sufficient to constitute a violation of the preventive arm, but it is not automatic if only one is breached.

Overall the spring forecast is the most importance data for the preventive arm as it is intended to assess an SDP might be launched and thus motivates the member state in question to take effective action to correct the violation for the current and following budget.

2.4. The Fiscal Compact

In order to further strengthen the preventive arm, the euro area member states (joined voluntarily by all non-EA member states except the Czech Republic and the UK) committed themselves to transpose the notion of MTOs into their respective national legal framework (ideally at constitutional level). This commitment is enshrined in the TSCG, which came into force in 2013. Its fiscal requirements are often referred to as the Fiscal Compact. It requires member states to achieve a structural balance no lower than -1% if the debt is below 60%. Otherwise the structural target is set at a maximum deficit of 0.5% GDP. An independent fiscal authority (e.g. a Fiscal Council) should monitor abidance of the rule and produce (or at least

---

[2] For member states, their MTO expenditures should not grow faster than the 10-year average growth rate of potential GDP. If the MTO has not been achieved, expenditure should go slower.

endorse) forecasts underlying the SCPs (based on a commonly agreed methodology). (European Commission, 2012c)

Moreover, an automatic corrective mechanism should be triggered in the event of a violation. If the member state is under an EDP, it should submit an Economic Partnership Programme, containing vital structural reforms it will undertake to improve long-term fiscal sustainability.

2.5. The corrective arm of the SGP

The corrective part of the SGP entails the famous 3% headline deficit threshold and the 60% debt-level target (both relative to GDP). The latter requires most countries to reduce their debt level at a 'satisfactory' pace. According the SGP regulation (EC-1467/97), the pace of the debt reduction is determined to be 'satisfactory' if "the differential with respect to the reference value has decreased over the previous three years at an average rate of one twentieth per year as a benchmark" (European Council, 2011). The deficit criterion is not met if the past, current or next year's deficit is significantly above 3% and the deviation is not merely temporary (European Commission, 2016a).

For the debt criterion, there are two phases that are evaluated against the targeted improvement and if none of them is complied with, an EDP may be launched:

1. A backward-looking assessment on the debt-level adjustment based on the year t relative to the three years before:

$$b_t > bb_t = 60\% + (\frac{0.95}{3})(b_{t-1} - 60\%) + (0.95^2/3)(b_{t-2} - 60\%) + 0.95^3/3(b_{t-3} - 60\%)$$

Where b_t is the actual debt ratio and bb_t is the derived benchmark debt ratio. (European Commission, 2016a)

2. A forward-looking assessment based on the debt level t+2 relative to t-1, t and t+1:

$$b_{t+2} > bb_{t+2} = 60\% + (\frac{0.95}{3})(b_{t+1} - 60\%) + (0.95^2/3)(b_t - 60\%) + 0.95^3/3(b_{t-1} - 60\%)$$

It should be noted that the Commission takes the year *t* to be the first outturn year and not the year in which the assessment is made.[3] Moreover a failure should not be attributed to the movement of cycle. The adjustment of the debt level to the cycle corrects the deficits of the past three years (CAB) and thereby recalculates the debt stock. Secondly, GDP (denominator) is adjusted as if the country had reached its potential growth rate over the past three years. It suffices if either of the non-cyclically adjusted debt ratios reached the target or that the cyclically adjusted target was achieved.

If a member state is in breach in the preceding year (t-1), an EDP may be launched and a recommendation encompassing a nominal and structural target for the in-year budget and following year's (t+1) balance is submitted to the member state. If deemed prudent, a later deadline for making the adjustment can be fixed.

Both deficit and debt criteria, if violated, require a minimum annual structural adjustment of 0.5%, but it can be set at a more ambitious level if necessary to reach the nominal deficit target. The Commission will also provide guidance on which action could help to achieve the targets (CSR). The first assessment is made ex-ante for the in-year and the following year, i.e. whether the plans are expected to lead to the desired structural improvement. If this is the case, the EDP will be put in abeyance. Once outturn data become available in the following spring, the EDP is

---

[3] "If the Member State is being considered for an EDP on the basis of its outturn data, the year t in the formula applies to the year that has just ended" (EC 2016a, p51).

abrogated if the targets were met and the then forward-looking assessment is also in line with the rules. If the targets are not met, a careful analysis will determine if 'effective action' has nonetheless been taken. The careful analysis should establish if the failure can be attributed to forecast errors or to an unanticipated, rapidly worsening economic environment. This analysis consists of a bottom-up approach (combined estimated effect of fiscal measures on the budget) and a top-down approach. The latter modifies the current estimates of the improvement in the structural balance for the influence of revisions. The revisions include those for the potential GDP, expected (windfall/shortfall) revenues and unexpected events. If either the bottom-up or top-down approach reveals that the government had intended to meet the targets, the procedure will be put in abeyance; otherwise, the EDP can be stepped up. (European Commission, 2016a)

If the Commission, in consent with the Council, asserts that the EA member states did not take effective actions, it can fine the member state a sum with a default value of 0.2% of the previous year's GDP, which will be transferred to the European Stability Mechanism (ESM). However, as shown in the case for Spain and Portugal on 27 July 2016, the Commission may suggest revise the fine to a lower value – in this instance even reduce it to zero (European Commission, 2016c). The motivation for this step has to be given publically, e.g. exceptional economic circumstances or an accepted 'reasoned request'. The latter must be submitted by the country in question 10 days following the decision by the EU finance ministers that effective action has not been taken.

If a fine is given and the violation persists in the following year (still without effective action being taken), the fine can be raised by a flexible value, but the total annual fine may not exceed 0.5% of GDP. Additionally, the violating member state (with the exception of the UK) faces the loss of access to the European Structural and Investment (ESI) funds and the Commission may invite the EIB to cut lending to the EA member state in question.

Overall both the EDP and SDP are launched based on data from the previous year or previous two years, while the forward-looking assessment is taken into account and can mitigate the former. The most crucial assessment is therefore conducted in April when the outturn data become available, although the autumn assessment can lead to a stepping up of the procedure as well.

3. Literature review on real-time fiscal data analysis

In an economic environment of sharply rising debt-to-GDP ratios, the European fiscal framework tools are under particular scrutiny as to whether they are capable of bringing the debt trajectory back on a sustainable path while allowing fiscal policy to act counter-cyclically. This has led to a stronger focus on structural variables but also raises the question of whether structural budget balances can be reliably estimated with the 'preliminary' and incomplete data available in real-time. Real-time[4] data analysis has been employed to highlight the uncertainties surrounding the output gap and CAB estimations.

In the past, real-time data analysis was predominantly used in the monetary policy domain (see Orphanides (2001), Croushore and Stark (2001), Orphanides and van Norden (2002 and Croushore (2011)). Similar research on fiscal policy is relatively new. One reason is the lack of adequate data, since fiscal variables (in a harmonised and persistent form) are published less

---

[4] This paper gives a broad definition to real-time data. "In-year" data contain estimates made in the year in question for the year in question. Forecasts are estimates for the year in question made in the years preceding it. Budget plans are forecasts of the preceding year (t-1). Outturn data or "ex-post" data are estimates made after the year in question. The 't+1' estimate is the first outturn data for a year 't'.

frequently than monetary variables and, even when available, do not date back far enough to allow for a robust analysis.

In the following sections, this paper will provide a literature overview of the main strands of real-time fiscal data analysis divided broadly into two groups with respect to their research aim: 1) statistical properties of revisions, and 2) the impact of political factors on revisions. Since this paper emphasises the first point, we will further break this section down into methodological approaches, the foremost being the chosen variable of interest.

3.1. Available data sources

One key challenge for real-time data analysis is the limited choice of datasets with a sufficient coverage of years and countries. There are various (national) datasets that provide most of the fiscal variables already in a useable format. For example, Beetsma et al. (2010) use Dutch data for their analysis, while Croushore and van Norden (2014) extract their data from the Federal Reserve Economic Data (FRED). However, in the European context, one would require harmonised data on several member states.

Researchers have started to collect data from each publication year and build their own real-time fiscal datasets by taking data from online databases or reports such as the spring and autumn forecasts or the SCP reports.

The choice of dataset depends on the research question and specifically on the necessary variables. The OECD data set is the most widely used one; see for example Golinelli and Momigliano (2006) looking at the primary balance, Cimadomo (2007) and Cimadomo (2012, update) using the CAB and Bernoth et al. (2015) focusing on the CAPB. The advantage of this data source is the fact that it offers the longest period of observations for the structural fiscal variables (around 20 years). Furthermore Hughes-Hallett et al. (2011) emphasise that the OECD's estimates are thought to be 'neutral' and less likely to be influenced by member states' 'political agenda'.

The IMF's World Economic Outlook (WEO), on the other hand, has its strength in the number of countries. Ley and Misch (2014) conduct their analysis on output gap data from the IMF and profit from observations for 175 countries.

More recently authors have turned to European data sources also for structural fiscal variables, such as De Castro et al. (2013) using Eurostat or as Beetsma et al. (2009), Frankel (2011) and Holm-Hadulla et al. (2010) using data published in the SCP.

Since the cyclically adjusted balance is the variable with the least observations, some authors such as Ley and Misch (2014) have calculated these based on the available information on GDP growth, elasticities, etc., e.g. via the HP filter method. However these constructed time-series will not 100% match those used by the policy-makers in question.

3.2. Magnitude of revisions

All studies dealing with real-time fiscal data firstly state some statistical properties of the revisions they obtain from their dataset, although not all of them focus on these.

The largest share of studies deal with the headline deficit or primary deficit, but more recently CAB analyses have become more numerous. Since statistical properties differ greatly according to which variable of interest is chosen, we split the findings into three sub-sections: headline budgetary balance, output gap and cyclically adjusted balance.

3.2.1. Headline budgetary balance

Researchers have found that budget plans are overly-optimistic relative to the outturn value, as shown by Marinheiro (2011), Frankel and Schreger (2013), Jong-A-Pin et al. (2012) and Beetsma et al. (2013) and others. The over-optimism is found across all EU countries although magnitudes differ (e.g. Gordo-Mora and Nogueira-Martins, 2007). In the same vein Frankel and Schreger (2013) and Paloviita and Ikonen (2016) find that policy-makers predict large deficits to be of a temporary nature whereas large surpluses are not predicted to decrease abruptly in the following year but only in the year after.

The magnitude of revisions for the headline budget balance is found to be significant for EU countries. Most studies find a forecast error in budget plans of 0.5%-1.0% of GDP, depending on the selection of countries, timespan and prediction time. Beetsma et al. (2013) obtained a mean absolute revision of budget balances of 0.5% for the EU countries (sample from 1999-2008). The analysis by Frankel and Schreger (2013) of data from 17 European countries based on the years 1995–2008 find a similar error of 0.52%. In the SCP, member states forecast the budget and GDP up to three years. Frankel and Schreger (2013) show that the revisions increase with the forecasting time horizon to 1.29% for the two years ahead and to 2.4% for a three-year ahead prediction.

The magnitude and causes hence vary depending on the vintage however De Castro et al. (2013) suggest that the systematic bias to downplay the deficit persists for four years (following the planning phase), but that the revision shrinks after two years.

It is important to note that the revisions in the headline deficit as a percent of GDP are not only driven by revisions in revenues and expenditures but also by the change in the denominator. Jonung and Larch (2006) find governments on average expect higher GDP growth than is actually realised.

3.2.2. Output gap

On average, studies on the EU conclude that the output gap is ex post estimated to be more positive than envisaged ex ante. For example, revisions for the EA between 2002 and 2010 calculated by Massimiliano and Musso (2011) were almost exclusively upwards. However since this time period encompasses the (largely undetected) bubble that was building up to the financial crisis but not many of the post-crisis years, our study may differ in this regard. Kempkes (2014) uses both Commission and IMF output gap data for EU15 countries between 2000 and 2012 and finds an upward correction between the autumn real-time and final (2013) of around 1 percentage point of GDP.

More importantly, studies' estimates on the magnitude of output gap revisions between the one-year ahead real-time forecast and final estimate (in absolute value) range from 0.5% to 2% of potential GDP, while most find a value of 1.0%-1.5%. Exemplary in a sample of OECD countries, Bouis et al. (2012) estimate that revisions of output gaps vary from 1% to 1.5% percentage points depending on specifications of country inclusion.

Similarly based on OECD data, Hughes-Hallett et al. (2011), looking only at EA member states between 1995 and 2008, find a root-mean-square error in the output gap of 1.25% average between the in-year and final estimate. Massimiliano and Musso (2011) highlight that the revisions of EA output gaps (1999-2010) are often as large as the gap itself. In line with this assertion, Tereanu et al., (2014) limit the revision period to what they refer to as the 'budget horizon', i.e. the forecast made in t-1 (fall) and t+1 (spring). They detect a mean absolute revision of 1.3% of GDP.

Kempkes (2014) compares the estimation performance between international institutions, namely the IMF, OECD and the European Commission. Looking at the difference between autumn t-1 forecasts and final value for EU15, he finds for the OECD data the (unweighted) average revision reaches a 1.9 percentage point negative bias (final output gap is larger than real time), higher than the 1.5 percentage points found for the IMF.[5] The European Commission's revisions are less pronounced with only 1 percentage point revision. Massimiliano and Musso (2011) uncover a high correlation between output gap estimates from different institutional sources. Several authors, notably Kuusi (2015) as well as Odor and Kucserová (2014a) have proposed alternative measures of the output gap and the structural balance which could be an added-value in the Commission's evaluation.

3.2.3. Cyclically adjusted balance

As summarised in the previous section, the headline budget balance tends to be more positive ex ante than ex post and the output gap turns out to be more positive than anticipated – both factors would lead to a worsening of the CAB[6]. Cimadomo (2007), based on OECD structural balance figures, shows a negative correlation between the output gap revision and CAB revision, i.e. a more negative output gap leads to a higher CAB.

The analysis by Kempkes (2014) emphasises the magnitude of cyclical component corrections (between the autumn t-1 forecast[7] and the final) equal to 0.5 - 0.9 percentage points downward for the EU member states on average. The mirror is a upwards revision in the structural balance of a similar magnitude; thus at least half a percent of potential GDP was 'spent' more than intended in terms of the cyclically adjusted measure. Tereanu et al. (2014), on the other hand, observe mean absolute revisions of the cyclically adjusted primary balance (CAPB) of 1.75 percent of GDP between t-1 (fall) to t+1 (spring). Hallett at al. (2011), looking only at euro area member states between 1995 and 2008, find a root-mean-square error in the CAPB of around 1.5% of GDP on average between the in-year and final estimate and 1.2% for t+1 vs. final estimate.

Ley and Misch (2014) used IMF output gap data for 175 countries from which they derive structural balances.[8] In their simulation the structural balance errors for the 10th and 90th percentile are at around 1.2% GDP.

Once again, the range of estimation errors greatly depend on the chosen estimation point in time (t-1, t or t+1) but the data source also drives the magnitude of revisions. Virkola (2014) shows that CAB forecasts made by the European Commission and the IMF are significantly different in real time with a gap of 0.3-0.4% potential GDP, although the estimates converge over time. The European Commission seems to be more pessimistic in real time on the output gap (more negative). Consequently the European Commission also sees a more positive CAB.

Overall the statistical properties of the CAB have been less thorough and detailed as for the other two measures. In section 5 this paper will provide a broader overview on magnitudes as well as dynamics, attempt to close knowledge gaps and verify assumptions with regard to the CAB.

---

[5] The OECD series has been used in its entirety; thus this average is based on more observations, which renders the figures less comparable.

[6] However, a revision of the output gap stems from a shift in potential GDP; thus a denominator effect may in some instances offset the impact and in others it may amplify it.

[7] The forecast made in t-1 for the year t.

[8] However, they make strong assumptions; among others are autumn, in-year forecasts for revenues and expenditures 'final' when projections are made for the following year.

Fiscal reaction functions using real-time CAB as the dependent variable indicate that for developed economies fiscal policy has been largely pro-cyclical or a-cyclical at best (see e.g. Cimadomo, 2012 and Alcidi et al., 2016). In connection with this observation, real-time data analysis has contributed by showcasing the uncertainty surrounding the cyclically adjusted balance. Most studies aim to show that the fiscal stance based on real-time data was more counter-cyclical than the ex-post outcome.

Forni and Momigliano (2005) were among the first to employ real-time data in fiscal reaction functions. They opted for a real-time output gap variable while maintaining the others in ex-post values. Looking at the structural balance as the dependent variable, they find counter-cyclical policy for 19 OECD countries for economic downturns.[9] This lends support to the notion that governments intend to act counter-cyclically, given their in-time knowledge of the cyclical position. Golinelli and Momigliano (2006) complemented the approach by additional control variables and concluded that the relationship also holds for economic boom phases.

3.3. Main findings and knowledge gaps

Recent academic literature using real time fiscal variables clearly show that revisions to the CAB, let alone the output gap itself, are large and significant – usually between 0.5% and 1.0% GDP. Several authors have shown a bias (on average) for the EU, which leads to an upward revision of cyclically adjusted deficits. Furthermore, the mainstream view is that governments act more counter-cyclically in their fiscal policy ex-ante than the ex-post data reveal. On the basis of the European Commission's estimates, it stands out that output gaps tend to be underestimated.

Most estimates on the statistical properties of CAB revisions are based on OECD data, but for the European fiscal framework it only matters if the Commission's estimate performs well. This paper will therefore focus on AMECO data and use a longer time series with the newest publications until 2016.

While many authors have pointed out the large revisions in the CAB, it has not been researched in detail if these constitute a problem for the functioning of the SGP, specifically if the revisions are likely to lead to costly policy errors or even unwarranted sanctioning. With this in mind, we will highlight which errors truly matter for the SGP in this regard and if there are safeguards within the fiscal framework to mitigate the impact of revisions.

Furthermore, to our best knowledge, the relevance of the 'turns of the cycle' with regard to the magnitude of revisions has not been established formally. At the turn of the cycle, potential economic costs of pro-cyclicality are likely to be particularly large. If the revision clusters around the turn, then critics would gain further ammunition, although it would also mean that during normal times the CAB is more precisely measured than the average revisions show.

4. Data description

A dataset has been created for the analysis which entails several variables, including the CAB, SB, cyclical component, net lending, potential GDP, trend GDP, nominal GDP and output gap. The original data were extracted from the AMECO database. The real-time dataset can be freely accessed via the FIRSTRUN website http://www.firstrun.eu/.

The data allow us the produce a (well-balanced) panel set of real-time (in-year) estimates as well as e.g. forecast estimates made in the year t-1 (i.e. the year prior to the year in question).

---

[9] Based on OECD data, but the period of observations was limited to 1993-2003.

Using the final-year estimates, the dataset can be exploited to calculate the revisions between the broadly defined 'real-time' and the final (2016) estimate. One could also envisage a model set-up in which revisions are calculated as the difference between the real-time and two-year past real-time estimates. This keeps the phase of revisions consistent across all vintage years, however since we are interested in the revisions vis-à-vis the 'true' value, the final estimate provides the best proxy.

Since the Commission has only recently started to report on the variables we are particularly interested in, the years used in our analysis run from 2003 to 2016, although for the revisions analysis, we exclude the last two years since these have not yet undergone sufficient revisions and would unduly lower the mean revision. Furthermore, we only consider data from the EA12 plus Denmark, Sweden and the United Kingdom, since other countries have fewer observations, which would bias our analysis over time.

5. An overview of fiscal data revisions

In this section we provide an overview on the data we use in the subsequent econometric section. Before running the regressions it is useful to analyse the data in its raw form. Many desirable and innovative insights can be gained by simple depiction of the time-series or its statistical properties, which have not been explicitly revealed in previous papers or only on a far shorter period of observations.

5.1. Summary statistics and dynamics

In the real-time literature, the emphasis has been placed on prediction errors with regard to the net lending, output gap and CAB. We follow the standard depiction of these forecasts but include also the 'backcast' (ex-post estimates) to show how revisions dynamics take place over time. For the European fiscal framework, the 'backcast' estimates are crucial for the EDP and SDP sanctions.

As a first visualisation, Figure 2 depicts the CAB over time. Each line represents the estimates made in the year listed in the legend below. In this graph Spain was chosen as an example, although revisions take a similar form for most of the other EA member states.

Figure 2. Spanish CAB estimate by estimation year (autumn)
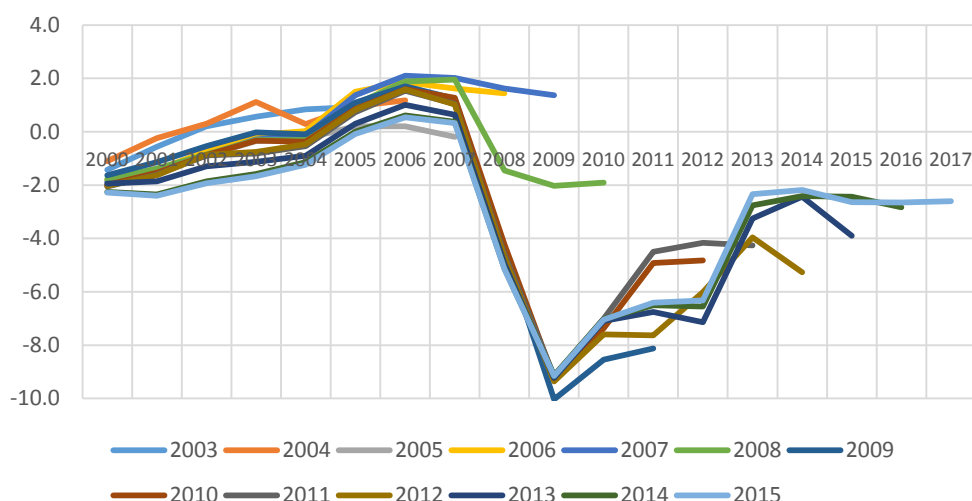
Source: Author's vizualisation on Commission data.

One can see that the forecasts are greatly revised, particularly during the crisis episode around 2008; however substantial revisions continued to take place in the following years. For Spain at the turn of the crisis the CAB was revised downwards 8 percentage points of potential GDP within one year. Real-time data indicate a long period of persistant cyclically-adjusted surpluses since 2001, although in hindsight there were only 3 years between 2000-15 in which Spain achieved a cyclically-adjusted surplus. It should be noted that revisions to the cyclically-adjusted fiscal stance were not one-directional but both overly-optimistic and overly-pessimistic at different times. Furthermore, adjustments are not limited to the in-year and forecast years, but extend to years that date more than five years back, owing, for the most part, not to headline budget revisions but rather to continuing output gap corrections, which we will formally show later on.

Looking solely at the gap between the real-time (in-year autumn) and the final estimates of the CAB, it is apparent that large revisions were observed for all EU countries in our dataset (see Appendix 2). Revisions for most countries peaked around 3% potential GDP, although Spain, Greece and Ireland reached values between 6% and 10%. For Ireland Appendix 3 shows a peak revision of around 20% of potential GDP in 2010[10]. In the following analysis, this particular country-year (IE-2010) will be excluded so as not to obfuscate the picture.

5.2. Revisions in the level of fiscal variables

Output gap

Table 1 illustrates the statistical properties of output gap revisions (in percent of potential GDP). The values were obtained over the entire panel of countries and over the years 2003-14. The revisions are calculated as the difference between CAB for year t made today (2016, i.e. 'final') and the CAB estimate for year t made in either t-1, t, t+1 or t+2. The table shows that accuracy improves the closer the time of estimation moves towards the final release. One might suspect that the European Commission over-corrects the output gap as a reaction to the adjustment in

---

[10] This revision is due to a reclassification of the bank bail-out; such a revision is not observed in the structural balance. This emphasises once more that ideally structural balance data would be preferable and should be used in future research once this data become available for a longer period.

the forecasting years, no over-correction is discernible, as implied by the shrinking mean absolute revisions.

Table 1. Revision of the output gap by estimation period

	Output gap						
	t-1 Spr	t-1 Aut	t Spr	t Aut	t+1 Spr	t+1 Aut	t+2 Spr
Mean	0.7	0.8	1.0	1.0	0.8	0.6	0.5
Stdev	2.5	2.1	1.7	1.5	1.4	1.1	1.0
Min	-6.4	-4.8	-3.8	-3.4	-3.2	-2.4	-2.4
Max	7.7	6.6	5.6	5.4	4.9	3.7	3.6
Median	1.1	0.9	0.9	0.6	0.6	0.5	0.3
Mean Abs	2.1	1.8	1.6	1.3	1.2	1.0	0.8

Note: With the exception of Ireland 2010, which is based on EU15, 'the revision for year t is caculated as the difference between the 'final' (2016) estimate and the estimate of listed publiaction time (e.g. spring t-1). The time span covered is 2003-2014.

The average revision, irrespective of the estimation period chosen, is positive and thus confirms the finding of previous studies that the output gap tends to be revised upwards in subsequent years. The mean revision for the in-year estimate is 1% of GDP, but it diminishes quickly thereafter to only 0.5% of GDP.

In essence the most illustrative gauge on accuracy is the absolute mean error which takes a value of around 2% of potential GDP in the spring forecast and declines to 0.8% for estimates made two years after the year in question. The mean absolute revision of almost 1% is still quite large and illustrates the fact that the output gap is constantly revised since the endpoints (on the forecast side) are changing and therefore shifting the potential GDP, which in turns alters retrospectively the output gap in years long passed. Overall, the impact of the output gap on the CAB will be negative since the mean of the output gap revision is positive, which we formally show later.

Headline deficit

Running the same exercise for the headline deficit (net lending) reveals a mean revision between -0.4% and 0% (see Table 2). Revisions for each set-up are negative; in other words over the whole period across countries the headline deficit was corrected upwards with respect to our final estimates and hence in line with the academic literature.

Table 2. Revision of the headline deficit by estimation period

	Net Lending						
	t-1 Spr	t-1 Aut	t Spr	t Aut	t+1 Spr	t+1 Aut	t+2 Spr
Mean	-0.3	-0.4	-0.1	0.0	-0.1	-0.1	-0.1
Stdev	3.1	2.6	2.0	1.5	1.1	1.0	0.8
Min	-13.1	-13.0	-10.0	-7.6	-5.1	-3.8	-3.7
Max	5.6	4.6	4.5	5.5	6.1	6.6	2.8
Median	0.4	0.1	0.1	0.1	0.0	0.0	0.0
Mean Abs	2.1	1.8	1.4	1.0	0.6	0.5	0.4

Note: With the exception of Ireland 2010, which is based on EU15, 'the revision for year t is caculated as the difference between the 'final' (2016) estimate and the estimate of listed publiaction time (e.g. spring t-1). The time span covered is 2003-2014.

The mean absolute error stood at 2.1% in the forecast year but subsides to 0.4% when the release date moves closer to the final estimate year. The extremes are far more pronounced on the downside, hinting at potential large revisions during the deep recession of the financial and euro crisis. These revisions are either due to a change in the budget in national currency or the denominator (nominal GDP). Since the headline revision is small but negative, it would lead to a deterioration in the CAB (if most of the improvement is not attributed to the cycle and thus only leading to an increase in the cyclical component).

Cyclically-adjusted deficit

Combining the summary statistics of the output gap and headline budget revisions, one would expect the CAB to be corrected downwards on average and indeed this is indeed the case (see Table 3). The CAB average level (not its revision) is estimated at around -1.5% and -2.0% over the predicting time.

The average revisions decreases from -0.8% to a difference of only -0.3% in the spring t+2 period vis-à-vis the final estimate. For all cases the balance is corrected downwards, which would lead to a surprise debt accumulation over the cycle (*ceteris paribus*). In terms of extreme revisions the downward shifts are three times as large as the upward shifts, (mainly due to the 2009 slump), but the median revisions remains stable at -0.4%. Mean absolute errors for forecasts are almost as high as the average CAB itself, but they decline to just above 1% for the in-year estimate.

Table 3. Revision of the CAB by estimation period

	Cyclically adjusted balance						
	t-1 Spr	t-1 Aut	t Spr	t Aut	t+1 Spr	t+1 Aut	t+2 Spr
Mean	-0.7	-0.8	-0.7	-0.5	-0.6	-0.5	-0.3
Stdev	2.7	2.4	2.1	1.6	1.3	1.0	1.2
Min	-13.0	-12.8	-10.4	-8.9	-6.0	-4.9	-4.4
Max	4.9	4.1	3.4	3.6	2.5	1.6	7.7
Median	-0.4	-0.4	-0.4	-0.4	-0.4	-0.3	-0.3
Mean Abs	1.8	1.7	1.4	1.2	1.0	0.8	0.8

Note: With the exception of Ireland 2010, which is based on EU15, 'the revision for year t is caculated as the difference between the 'final' (2016) estimate and the estimate of listed publiaction time (e.g. spring t-1). The time span covered is 2003-2014.

For the CSR the error of 1% on average in spring t+1 combined with the even larger error for the forecasts could lead to (ex post) ill-fitting policy advice – whether such a deviation is welfare increasing or decreasing is not pursued in this paper, but it is merely acknowledged as an unintended fiscal stance. Nevertheless one can imagine that a recommendation to tighten fiscal policy might unnecessarily jeopardise economic growth. Conversely the lack of pressure to tighten policy, when ex post the rules would have demanded this, will affect debt sustainability and push the member state further away from its MTO, thus potentially leading to higher fiscal efforts required in subsequent periods.

It is worth noting that both headline and cyclically adjusted balances are prone to estimation errors. In fact the headline deficit revisions are nearly as large as for the CAB with regard to the forecast and in-year estimates and only marginally smaller for t+1 vintages.

5.3. Revisions in the structural improvement

For the European fiscal framework, the revisions in the level of the CAB are of importance to the MTO, but revisions in the CAB growth rate are just as pivotal, since targets are also given in the form of structural improvement. This stems from the requirements in the SGP to take adequate measures to rectify the headline deficit by achieving an improvement in the structural balance. Table 4 lists the growth rate of the variable estimated in 2016 minus the growth rate estimated in t-1, t or t+1.

The output gap mean (absolute) revision is much lower for this measure than for the level, since the output gap revision usually lifts potential output. The data show a strong improvement in accuracy for the t+1 autumn backcast, both for the change in the output gap and net lending. Interestingly net lending does not have an upward nor downward bias, whereas the output gap still has a small upward bias. The mean absolute revision remains high at 1.6% and 1.4% for forecasts, but in autumn t+1 subsides to only 0.5% in both cases. Overall growth revisions tend to be somewhat smaller than the revisions in levels.

Table 4. Revisions of the change in annual output gaps and net lending

	Output Gap						Net Lending					
	t-1 Spr	t-1 Aut	t Spr	t Aut	t+1 Spr	t+1 Aut	t-1 Spr	t-1 Aut	t Spr	t Aut	t+1 Spr	t+1 Aut
Mean	-0.4	-0.1	0.2	0.3	0.3	0.2	-0.1	-0.2	0.1	0.0	0.0	0.0
Stdev	2.2	1.8	1.3	1.2	0.9	0.7	3.1	2.0	1.7	1.9	0.9	0.7
Min	-8.2	-7.5	-3.5	-3.7	-2.5	-1.9	-14	-6.1	-9.8	-12.8	-3.1	-3.2
Max	5.5	4.5	4.1	4.1	3.9	3.4	7.7	3.6	3.5	4.2	3.7	3.8
Median	-0.2	-0.1	0.1	0.4	0.2	0.2	0.2	0.1	0.2	0.0	0.0	0.0
Mean abs	1.6	1.4	1.0	0.9	0.7	0.5	2.1	1.4	1.2	1.3	0.5	0.4

Note: Excluding Ireland 2009-10, based on EU15.

Given the low mean revisions for the output gap and net lending, it is not surprising that also the mean error of cyclically-adjusted balance improvement is only marginal and centred around zero. The magnitude of revisions stays at just above 1% for the forecast and in-year estimates but drops sharply to 0.5% for the t+1 estimates. The Commission's estimates of the change in the CAB are more precise than the level predictions. It is likely that the absolute errors would have been even smaller if the peak years had been excluded, since the peak year's CAB will be corrected downwards (more positive output gap) and the first 'bust' year CAB will be shifted upwards (more negative output gap).

Table 5. Revisions of the change in annual CAB

	Cyclically adjusted balance					
	t-1 Spr	t-1 Aut	t Spr	t Aut	t+1 Spr	t+1 Aut
Mean	0.1	-0.2	0.0	-0.2	0.0	-0.1
Stdev	1.8	1.6	1.7	1.8	1.0	0.8
Min	-5.3	-5.7	-10.4	-11.4	-3.0	-3.2
Max	7.6	4.4	3.5	3.7	4.4	4.2
Median	0.1	0.0	0.2	-0.1	-0.1	0.0
Mean abs	1.3	1.1	1.1	1.2	0.6	0.5

*Note*: Excluding Ireland 2009-10, based on EU15.

The summary statistics have already shown that the mean as well as the mean absolute error decrease the closer the estimation date of the data comes to the final estimation year. Appendix

3 shows support for the claim of Cimadomo (2012, p. 30) that the forecasted CABs are "bias estimates for their final values". Both Cimadomo (2012) and Croushore (2011) argue that revisions reduce 'noise' instead of adding news. Appendix 3 shows visually how the accuracy improves over time (i.e. moving closer to the final estimation year). Revisions between the period t+2 and the final estimate were in around 30% of the country-year estimates larger than 1% of potential GDP, 16% were revised by more than 1.5% of potential GDP and only 9% were corrected by more than 2% of potential GDP. This hints once more that the uncertainty of output gap estimation remains even two years after the year of 'observation' has passed. It thus worthwhile to give a quick glance at the correlation between the output gap revision and CAB revisions.

5.4. Causes of revisions and particularities

Output gap revisions are clearly one of the key drivers of the corrections in the CAB but for the adjustments made between the t-1 and the final there is no significant correlation visible (see Figure 3), in other words, here the revision still originate in changes to the headline budget or nominal GDP. For the in-year correlation we obtain a clearer relationship with a negative slope. The coefficient does not change when moving to a t+1 backcast but the goodness of fit rises to 38%.

Figure 3. Correlation between output gap revisions vs. CAB revisions, 2004-2014
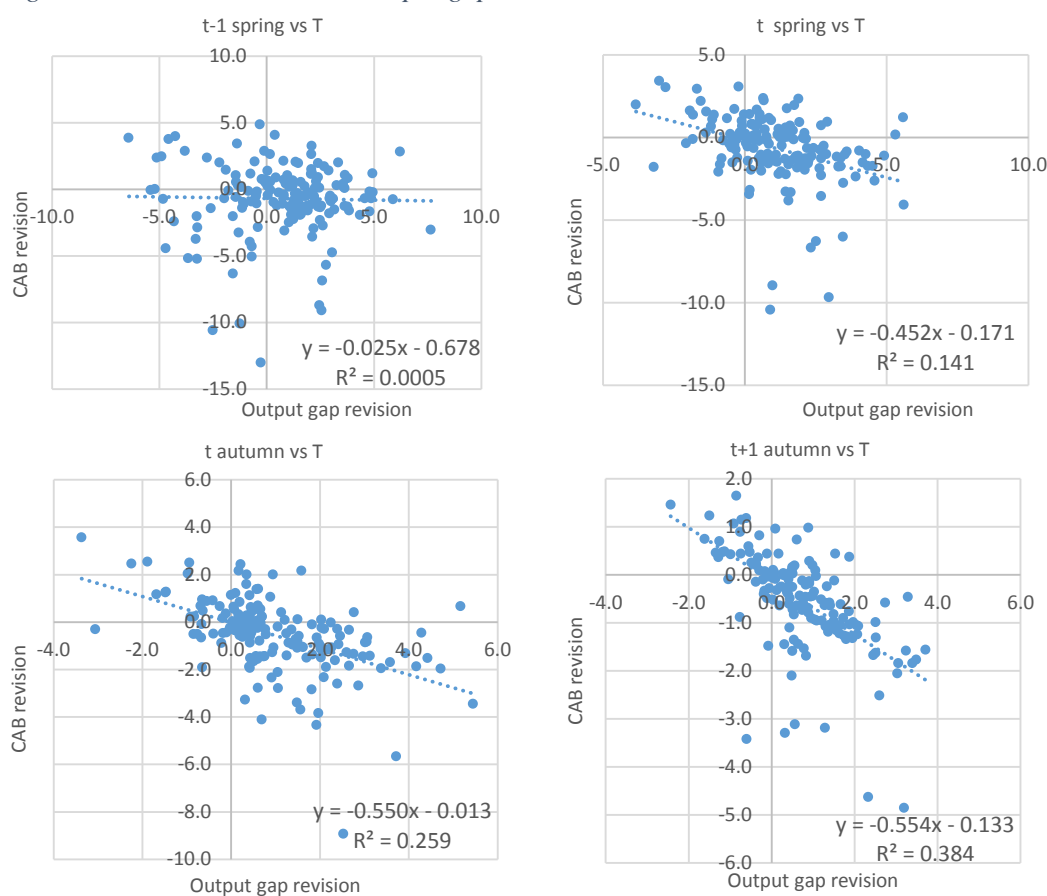

Note: Calculations exclude the country-year estimate for Ireland-2010; otherwise all years for all countries are displayed.

Since the output gap is clearly one of the key drivers (at least in the latter estimation years), a look closer at this variable's revisions is merited.

Figure 4 shows the real-time output gap on the horizontal axis and the final output gap on the vertical axis. Similar to the CAB, the dispersion around the 45 degree line is in most instances large. Most output gaps cluster around -2% for the real time and 0% ex-post. Revisions tend to be particularly large and positive for the years leading up to the financial crisis (2006, 2007 and 2008). The downward revisions were less pronounced but for the most part centred around 2009 and 2012, just when the double-dip recession ensued.

Figure 4. Correlation final and real-time output gap by year, 2003-15
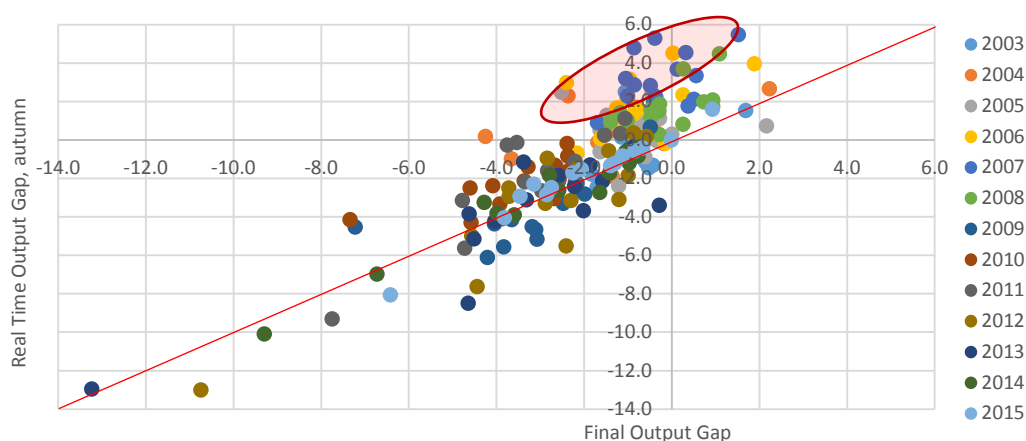
Switching to an analysis by country instead of by year one has to note that the data points on the far lower left and upper right are the figures from Greece. Without them, the trend line would be substantially steeper than the 45 degree line. If the output gap in real time is positive, the revision is almost exclusively revised upwards, this shows the underestimation of the boom phase in real time. Also small negative output gaps (below 2%) in real time are revised upwards on average with a large share being revised to a positive output gap. For large negative output gaps, the relationship turns (strongly driven by Greece and Spain). Overall the output gaps are underestimated, i.e. that are predicted to be lower than outturn data reveal. Belgium and with one marginal exception France, Italy and Portugal revised their output gap estimation upwards for all years, thus exhibiting a clear bias to see the output gap more negative than it actually is.

The summary statistics and graphs have already provided a good overview on fiscal data revisions and their properties. They have verified important findings of other authors with more recent data and are crucially based on data that are relevant for the EU's fiscal framework. Figure 4 hints at the answer to one of our main questions: Are revisions of the cyclically-adjusted balance larger around the turn of the cycle?

6. Analysis: Relevance of the cyclical position

Our methodological approach is first to calculate several statistical properties for revisions vis-à-vis a particular time of measurement. We furthermore use the raw data to evaluate in how many instances corrections in the variable defined as the first difference in the cyclically adjusted balance would have shifted it above or below the 0.5% minimum structural

improvement requirement. Complementarily we show whether corrections in two consecutive years[11] moved in the same or opposing directions.

The main ambition of this section is to showcase the reliability of real-time fiscal data around the turning point of the cycle, i.e. to show whether revisions are particularly large for those years. If the magnitude of revisions are largely owing to estimation errors around the turns, the warning and correction system that is the SGP would be least effective when it is needed most, although in normal times the revision would therefore be smaller and more reliable than indicated by the average revision[12]. If the turn of the cycle does not significantly influence the estimation performance, then the corrections would be systematic but also smoother.

We ran a random-effects[13] GLS regression of the CAB revision as the dependent variable and the size of the output gap as the independent variable:

$$(CAB_t^T - CAB_t^t) \ = \alpha + \ \beta OG_t^T$$

with CAB_t^T representing the cyclically adjusted balance for the year t made in the final outturn year[14] T and CAB_t^t being the real-time estimate. The independent variable is the final output gap. The correlation exercise is repeated with various forms of 'real-time' data, i.e. estimations made in t-1 and in-year t and t+2.

Table 6. Correlation of CAB revisions and output gap

	t-1 spring	t-1 autumn	t spring	t autumn	t+1 spring	t+1 autumn
Final output gap	-0.179**	-0.153**	-0.239***	-0.264***	-0.262***	-0.185***
	(0.06)	(0.05)	(0.04)	(0.03)	(0.02)	(0.02)
Constant	-0.855**	-0.966**	-0.984**	-0.874***	-0.913***	-0.768***
	(0.31)	(0.32)	(0.31)	(0.23)	(0.17)	(0.13)

* p<0.05, ** *<0.01, *** p<0.001

Indeed the output gap and the CAB corrections are correlated in the panel regression for all specifications. The null-hypothesis is rejected at 1%[15] and the 'within R^2' reaches 50% for t+1 estimates. The output gap coefficients are negative throughout estimation periods. In other words, given the output gap coefficient of around -0.25, a 1% higher output gap (measured ex-post) is associated with a smaller (or rather more negative) revision in the CAB on the order of 0.25% of potential GDP.

This regression is step-wise complemented by dummy variables for the turn of the cycle, peak, trough and interaction terms, in order to evaluate if these impact the magnitude of revisions:

$$(CAB_t^T - CAB_t^t) \ = \alpha + \ \beta_1 (OG_t^T) + \ \beta_2 (OG_t^T) * Dummy_t^{turn} + Dummy_t^{turn}$$

where the superscript denotes the time of estimation and the subscript the year for which the estimate was made. For the dummy the superscript states which variable the dummy represents. The regression is applied to revisions between t-1, t, t+1 and t+2 (for both spring and autumn) vis-à-vis the final estimates of spring 2016.

---

[11] For our purposes both are based on outturn data for t-1 and t-2.

[12] Assuming turns are far less frequent than 'normal times'.

[13] A Hausman test reveals that the random-effects model is appropriate and the difference to the fixed-effects model coefficients is not systematic.

[14] Spring 2016.

[15] 0.01 for the forecast years t-1.

The first challenge in the analysis is to determine the turns in the dataset. Establishing the 'turn of the cycle' in essence boils down to a switch in the sign of the annual change in the output gap. If the gap to potential output is positive, a sudden downward shift signifies a turn from a widening gap to a closing gap. Hence in a first step the years are selected in which the growth rate of the output gap changed signs. The data used are final estimates of the output gap.

For the purpose of the paper, the 'turn' dummy consequently does not signify the actual peak and trough of the cycle but the year just after these, since in that year the view on the direction of the gap changed.[16] However not every change in the sign truly constitutes a 'turn' of the cycle but may merely be attributed to the small volatility around the trend. In order to discern the actual turns, one can use a minimum threshold in the change of the output gap for years when the sign switches. In order to avoid counting very small fluctuations (ups and downs) as a turn, we establish a minimum change in the output gap of 1% potential GDP between these two years. Choosing such a threshold is by default arbitrary but graphical analysis revealed that excluding the turns below 1% results in a reasonable interval of peaks and troughs.

Table 7 shows the results of the correlation with includes either a dummy for the turn or dummies for the peak and trough. The previously determined negative relationship between the output gap and the CAB revisions still holds. The coefficient for the turn dummy is as expected negative, i.e. it amplifies the relationship between output gap and CAB revision. However, the turn dummy is not significant.

In order to verify the robustness of this results we conducted a sensitivity analysis detailing the impact of splitting the turn into peak and trough, excluding Ireland due to its outlier character in terms of corrections and lastly verifying whether the relationships hold for the in-year as well as the t+1 estimates.

Table 7: Correlation overview by specification

	GLS Regression results			
	in year estimate (s)		t+1 estimate (s)	
	M1	M2	M1	M2
Final output gap	-0.177**	-0.168**	-0.266***	-0.281***
	(0.06)	(0.06)	(0.02)	(0.02)
Turn dummy	-0.51		-0.086	
	(0.36)		(0.13)	
Peak dummy		-0.778		0.373*
		(0.49)		(0.17)
Trough dummy		-0.402		-0.375*
		(0.44)		(0.15)
Constant	-0.815**	-0.789**	-0.836***	-0.860***
	(0.34)	(0.34)	(0.17)	(0.17)
Final output gap (absolute value)	0.131	0.120	0.013	0.008
	(0.08)	(0.08)	(0.04)	(0.04)
Turn dummy	0.784*		0.005	
	(0.32)		(0.15)	
Peak dummy		0.760		-0.064
		(0.43)		(0.20)

---

[16] If the model is estimated using the actual peak and trough, all set-ups result in insignificant correlations.

Trough dummy		0.736		0.098
		(0.38)		(0.18)
Constant	0.968**	1.01**	0.961***	0.965***
	(0.32)	(0.33)	(0.14)	(0.14)

Notes: Fixed-effects regressions; interaction terms not displayed; value of the output gap estimate is from 2016. The lower panel uses absolute values for the output gap and the cyclically adjusted budget balance revision. The revisions are calculated as the difference between the in-year estimate and the final year (column one) and the t+1 estimate and the final year (column two). *, **, and *** denote, respectively, significance at the 10, 5 and 1 per cent level. Standard deviation in parentheses.

The turn is not significant for the first model which includes Ireland and even when the outlier Ireland is excluded or one moves to the t+1 estimate the existence of a turn still does not matter. Separating the peak and trough shows no significant relationship between the revision and peak or trough. Only for t+1 estimates can we find any (marginally) significant results. The peak and trough coefficients are of the same magnitude only that the former is positive while the latter is negative. The signs are contradicting the prior expectation since one would assume the CAB to be revised downwards (worse than anticipated) after the peak.

A graphical analysis depicting the relationship between the absolute value revisions and absolute value output gap suggested that there may be a positive correlation between the two. However, our results in the bottom panel of the table above show that this relationship is not significant. Overall the evidence that CAB estimates are particularly large at the turn of the cycle are relatively weak.

The insights obtained from the previous section and the regression results in this section lead to the conclusion that the revisions of the CAB may be centred on the peaks and troughs of large cycles, however this is not the case for peaks and troughs of smaller cycles and thus the overall relationship is less clear. One reason may be that the revelation of a sudden economic recession triggers a revision not only for the next years but just as importantly for the previous years. Large revisions for year preceding the turn of the cycle are largely caused by revision of the potential GDP and thus the gap itself. Thus revisions seem persistent over the entire cycle and may not mainly be foremost attributed to the turns.

As the fiscal framework relies not only on levels but moreover on structural 'improvements', the model is re-estimated with the change in the CAB. The delta CAB variables have been found to be less prone to revisions and since they encompass two years. The already-weak correlation for levels disappears for the prediction error in the annual change in the CAB. In other words, the structural improvement is mis-measured irrespective of whether the year signifies a turn of the cycle or not.

Using a scatterplot may show the results more illustratively (see Figure 5). Of course the deep crisis which began in 2008 exhibits higher revisions than the average, mostly due to three outlier countries (see Appendix 4). The year 2008 witnessed downwards revisions from all countries but also the boom years leading up to the crisis were marked by substantial revisions in some countries. The absolute revisions per year (for the in-year estimate in autumn) mostly range between 1% and 1.5% of GDP. The year 2008 revisions stood at 2.7% – clearly above average. The year 2009 was above the average with 2.1%, but this showing was mostly driven by the performance of Ireland and Greece. Over the entire panel, it is not the case that the peak or trough years appear to be more error-prone than normal years.

Figure 5. Revisions of the CAB by year
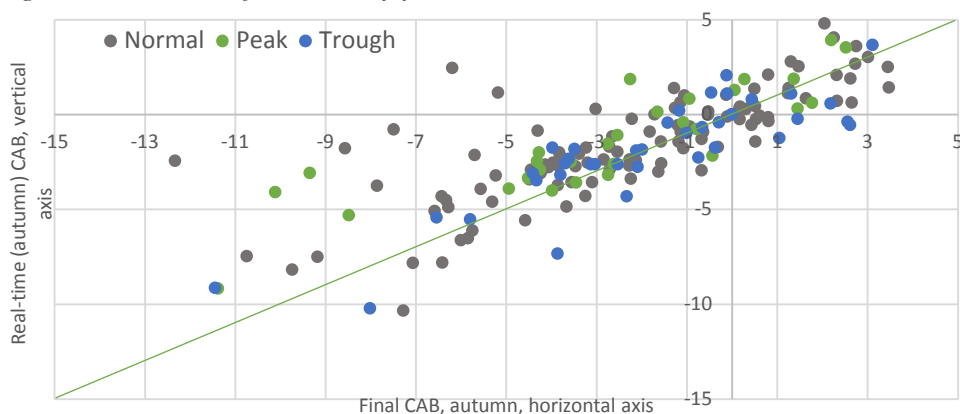
It is thus true that revisions were the largest at the height of the biggest cycle, but this is not the case for smaller turns before or after. One reason of course is the persistent revisions of the headline deficit and output gap, which are not limited to peaks or troughs. Secondly a large turn will emit revision waves leading to corrections in the phase leading up to the trough and in particular to the peak.

Of course the analysis would profit from a longer time period with a larger number of large boom-bust cycles and in future it would be beneficial to repeat the exercise not only with more observations (once they become available), but also by using various stricter definitions of a (large) turn, i.e. above 2% GDP.

Nevertheless, given the available data and our parameters, it seems that revisions are not as a rule larger at the turning point, but rather spread like waves backwards and forwards, thus 'smoothing' the magnitude of revisions across the entire period. If this is the case, the mean revisions are not due to one 'off-year', but rather are systematic. As a consequence, the question now arises: does this finding jeopardise the aims of the SGP, as is often suggested in policy debates?

7. Impact on the European Fiscal Framework

Larger revisions in the CAB have not been found to cluster at the turn of the cycle and are of a systematic nature, i.e. present at peaks, troughs and normal times. Hence the persistent, large revisions may lead to policy advice or sanctions given in real-time which are ill-fitting or unwarranted in retrospect. These errors could potentially be economically and politically costly.[17]

7.1 Unwarranted stepping up of the EDP and SDP

Assuming, for the sake of argument, that all countries in our sample had been required to achieve the established minimum structural improvement of 0.5% of GDP per year, as laid out in the SGP corrective arm for member states under EDP[18]. We can then derive the number of instances in which the ex-ante and ex-post assessment coincide or not. We calculate the

---

[17] Ex-post ill-fitting advice based on the rules may nonetheless lead to welfare gains instead of welfare losses. This depends on whether the welfare maximising structural stance is in line with the one established under the SGP – a question that is beyond the scope of this paper.

[18] As established in the SGP framework requiring a benchmark structural improvement at least 0.5% of GDP which should be line with correcting the nominal deficit to a value above the 3% headline deficit.

revisions of the ex-ante change in the CAB between year t-2 (spring) and t-1 (spring[19]) made in year t vis-à-vis the 'final' estimate of delta CAB for the same years. Our data sample encompasses the years 2004-14 and shows that only in 14 instances out of 163 observations would a sanction given in real-time have proven to be unwarranted ex post. The failure to give an ex-post deserved sanction would have occurred in 10 cases. Nearly 85% the advice would have been fitting ex ante as well as ex post. Next we only considers years after an EDP had been identified, i.e. those years a member state is asked for structural improvement and potentially a stepping up of the procedure could lead to a sanction. This reduces our sample to 64 observations but the success rate does not changed much. Only in 8% of the cases would a sanction have been unwarranted and in 8% missed. Expressed as a percentage of all years in the sample, only 4 in 100 country-years were the sanctioning decisions inappropriate.

As a refinement one can use the required structural improvement under the preventive arm. In the preventive arm, member states are required to achieve their MTO or 'sufficiently' adjust towards their MTO. The MTO requires most countries to achieve a (small) structural surplus since most member states exhibit debt-to-GDP ratios over 60%. In contrast to our example with the standard 0.5% structural improvement we now establish the required structural improvement according to the Commission's methodology under the preventive arm, which takes into account the position of the business cycle as well real and potential GDP growth rates (see Figure 6).

Figure 6. Matrix specifying the required adjustment towards the MTO

		Required annual fiscal adjustment*	
	Condition	Debt below 60 and no sustainability risk	Debt above 60 or sustainability risk
Exceptionally bad times	Real growth < 0 or output gap < -4	No adjustment needed	
Very bad times	-4 ≤ output gap < -3	0	0.25
Bad times	-3 ≤ output gap < -1.5	0 if growth below potential, 0.25 if growth above potential	0.25 if growth below potential, 0.5 if growth above potential
Normal times	-1.5 ≤ output gap < 1.5	0.5	> 0.5
Good times	output gap ≥ 1.5	> 0.5 if growth below potential, ≥ 0.75 if growth above potential	≥ 0.75 if growth below potential, ≥ 1 if growth above potential

Source: European Commission (2016a).

Out of the 165 observations, 141 assessments of compliance within the preventive arm would have been identical ex ante and ex post, i.e. 85% of the time. In only 7% of the cases would a stepping up of the preventive arm (in the form of a deposit) have been unwarranted, and a deposit requirement would have been overlooked in 7% of the cases. Thus the performance with regard to the preventive arm is similar to revisions under the corrective arm. Notably 1/3 of unwarranted sanctions would have been given to Luxembourg due to significant revisions in several years. Conversely France would have unduly avoided sanctioning the most. In reality no

---

[19] Since the reform of the SGP, the autumn evaluation for EA member states may also lead to a stepping-up of the excessive deficit procedure, but for our analysis only spring assessments are considered.

deposit has been asked for under the preventive arm so far since mitigating circumstances[20] were found (as elaborated in the next section).

If only those years after a country entered into an EDP are assessed, unwarranted sanctions would have only occurred in four cases (or 5% of the time), while missed sanctions were twice as likely. The reason being that selecting only EDP years focuses our sample on the crisis years. In this light it is remarkable that the share of incorrect assessment did not increase. The explanation is that the years of "exceptionally bad times" do not require any adjustment, and thus a revision would not be picked up in our simulation. Indeed in 62 observations no adjustment was needed, thus rendering revisions irrelevant for this exercise. In both simulations the rate of error is relatively small, given the magnitudes and systematic nature of the revisions as detailed before. Moreover, the simulation was conducted with a strict target and a small shift below the target (e.g. from 0.6% ex ante to 0.4% ex post would be classified as an incorrect assessment), whereas in reality some discretion is possible as described in the next section.

7.2 Safeguards and mitigating factors within the SGP

Firstly, the forecast-year revisions of the CAB are large, often as large as the variable itself, but in the first ex-post year the estimates are distinctly smaller in absolute value and the average is relatively close to zero. Forecasts feature prominently in the SGP, but they matter less in terms of putting pressure on member states. Projections do play a role in the forward-looking debt criterion, the assessment if a member state is on a credible adjustment path towards the MTO can determine if an EDP is put abeyance, but it is not the most decisive for sanctions. To be sure, ill-fitting policy advice may still rest on unreliable forecasts, but they will not (by itself) trigger fines. For these, outturn data are decisive, even if the ex-ante assessment is taken into consideration as well. The strongest pressure on member states thus stems from the assessment of the first ex-post data (t+1). In order to avoid a fine, a member state may take more drastic measures and the deviation in the outturn data from the target determines the amount of improvement demanded for the subsequent years.

Even in the preventive arm the launch of an SDP is based on outturn data – either a deviation from the adjustment path of 0.5% of GDP in the past year or 0.25% GDP as an average over the past two years. It is important to note that either violation can trigger an SDP. Moreover, since the adjustment required to reach the MTO within the deadline may shift due to the CAB correction of the previous year, the demanded structural improvement is frozen at the time of the assessment. This latter step prevents a member state from having to increase or decrease its structural efforts merely because revisions of the preceding year took place over the course of the current year. However, if the revisions lead to a member state already reaching the MTO in the previous year or if the output gap is revised in such a way that it would require a lower structural improvement (see Figure 1), then the revisions take precedent over the old assessment. This system ensures that the member state clearly understands how much structural effort is required, even if the levels are revised but at the same time ensures that large revision in favour of the country will be taken into account. Moreover, despite eventually triggering a sanction, the pressure under the SDP is weaker than in the EDP since the sanction can only be imposed in the form of an interest-bearing deposit of 0.2% GDP.

We turn now to the debt criterion in the corrective arm, in which a member state must show a debt reduction (over three years) of 1/20 of its access debt[21] either backwards-looking or

---

[20] This includes the fact that in reality the required improvement can be assessed over several years instead of a single year which may be in violation.

forward-looking. It suffices if the cyclically adjusted debt level has been reduced accordingly, i.e. by approximation a structural deficit target has been reached over the past three-year period. Ultimately revisions over the three years in question taken together determine whether the real-time assessment still holds with ex-post (years later). Since the cyclically-adjusted view depends on the three years leading up to the assessment year, we have analysed whether the revision of the year t-1 and t-2 (based on the estimate in year t vis-à-vis the final – i.e. 2016 – estimate) of the structural improvement (delta CAB) has moved in the same direction or whether the revision over two years tends to offset each other. Looking at the years 2005-14 reveals that revisions are slightly more likely to move in opposite directions (77/148). More importantly, in terms of amplifying or mitigating magnitudes, the opposing revisions of the year t-2 are 1.5 times larger than those going in the same direction. This suggests that taking a two- or three-year average lessens the impact of revisions and improves (ex-post) accuracy. The aforementioned SDP is launched if the one-year or two-year average targets are not met – it would have been beneficial if the rules had specified 'and' rather than 'or', in order to reap the same benefit here.

Nevertheless, both the debt criterion and the preventive arm are equipped with safeguards limiting the impact of revisions. However, what happens once the EDP has been launched (either on the basis of deficit or debt criteria)?

The member state is asked to achieve an annual minimum structural improvement of 0.5% of GDP. As shown above, the likelihood that the wrong policy advice/sanction is administered, given the historical data, is around 1/6 of the time. Our summary statistics show an absolute value of the CAB revision at 0.5% of GDP. Notably when assessing the fiscal effort,[22] the European Commission verifies its findings via their top-down approach, which checks whether the deviation from the target was not due to revisions in the potential growth rate. Hence, here also we find that the impact of revisions is mitigated by the rule specification.

Moreover even if a numerical violation is detected, it would not automatically lead to a sanction since a deviation of 0.25% from the structural improvement target is not considered a 'significant deviation'. Furthermore the European Commission can make use of several exception clauses (structural reforms, events outside of the member state's control, EA-wide downturn), to avoid establishing a 'significant deviation'. A 2015 communication from the European Commission, which was endorsed later by the Council, stressed the benefit of making full use of the flexibility already present in the SGP (European Commission, 2015). The recent example of Spain and Portugal highlights that even if the formal exemption clauses are not available, the Commission alongside the Council can still abstain from imposing fines on a member state if they find vindicating factors (European Commission, 2016c). Spain and Portugal did not achieve the required structural adjustment in 2015 but the Iberian countries submitted a 'reasoned request' not to impose a fine (European Commission, 2016c and 2016d). The Commission accepted the reasoning and recommended to the Council not to impose a fine despite numerical violation.

All these built-in flexibilities render (ex-post) unwarranted sanctions even more unlikely than the numerical probability (based on historical data) would suggest. It also reduces the pressure and thus the likelihood of ill-fitting advice to tighten fiscal policy.

---

[21] Difference to the target maximum debt level of 60%.

[22] I.e. improvement in the structural balance.

These safeguards, however, are asymmetric and given the average correction of the CAB downwards, it raises the question of debt sustainability. Kempkes' (2014) simulation based on European Commission data from the EU15 finds that the debt trajectory is altered by revisions on the order of 6 percentage points over a period of 10 years. He suggests using a control account to adjust the structural balance for the bias. This would also be politically difficult to adopt and 6% over 10 years is comparatively small. Moreover the Commission, in setting the MTO, has the ability to create a buffer that ensures debt sustainability in the long-run. For example, a deficit of 0.5% may be augmented to 1% if debt levels are high and the sustainability MTO will require the member state to achieve a balanced or surplus MTO. Thus on average the ex-post stance will not push the debt trajectory to unsustainable levels. Notably the MTOs will be rounded downwards to the nearest 0.25% of GDP (for example, from -1.05% to -1.25%). Perhaps given the average downward revisions, it should rather be always rounded up to provide an additional buffer – although it may create counterproductive incentives.

More broadly, at the core of the issue is the lack of savings during the boom phase. The relatively new expenditure benchmark which is now part of the preventive arm of the SGP, may prove effective in limiting the 'overspending' during good times. Despite the fact that also the expenditure benchmark relies on potential output growth estimates, it has been shown by Kuusi (2015) to improve counter-cyclicality. Secondly, wind-fall revenues are to be used for debt reduction, and if properly monitored, they may suffice to correct the debt accumulation due to revisions of the CAB.

Overall, the SGP is well-equipped to respond to revisions in the CAB, at least for average-sized revisions. It is noteworthy that the headline deficit revisions are for the in-year estimate nearly as large as for the CAB. The difference is that the CAB still shifts years after the budget has been finalised, but at that point revisions are relatively small and of less concern within the SGP. More importantly, a rule based on structural balance enables member states, despite significant revisions, to pursue a counter-cyclical fiscal policy more than any headline target would allow.

8. Conclusion

Cyclically-adjusted balances are exposed to large revisions over time, which are not limited to in-year estimates or forecasts, but also apply to data that are already classified as 'outturn' data. Forecasts made in the preceding spring are on average revised by 1.8% of potential GDP. In-year estimates of the cyclically-adjusted budget balance are revised on the order of 1.4% of GDP and the first outturn data by around 1% GDP. Our analysis reveals an average downward revision of around 0.5% GDP for the in-year estimate, i.e. the structural balance is worse than anticipated. This is largely owing to an upward revision of the output gap, in particular outturn data revisions stemming from the output gap corrections.

Our analysis of the pattern of revisions with regard to the business cycle reveals that the CAB revisions are not larger at the turn of the cycle. However, substantially larger (downward) revisions were triggered at the height of the crisis in 2008-09 than in any of the other years examined in our sample. The sudden revelation that 2008 indeed was a peak year and the fact that the ensuing deep economic recession triggered not only revisions for this period but that they also spread like waves back towards the pre-crisis years when, in retrospect, output gaps has been heavily underestimated.

Since as a rule the turns of the cycle do not signify years of particularly large revisions, the uncertainty is spread across all years and proves to be a systematic problem. In other words, the average magnitude of the revisions applies to most years.

This raises doubts over the viability of relying on structural variables for a numerical rule-based fiscal framework. For the Stability and Growth Pact the revisions of forecasts and in-year estimates may in some instances lead to ill-fitting policy advice, but the real pressure on member states to respond to the European Commission's demand originates from the outturn data, i.e. the determination of whether a sanction will be imposed. The average revisions of estimates for the year following the year in question are not as high, at only 0.8% of potential GDP. Moreover the CAB level is less decisive since the member states are, under the EDP, required to achieve a structural balance improvement over the preceding year or preceding two years. The change in the CAB is 'only' revised by 0.5% of potential GDP. Furthermore our analysis has revealed that revisions tend to be mitigated when looking at the improvement of the CAB over two years instead of a single year – i.e. a revision for one year offsetting the one for the other year.

The SGP also entails safeguards against revisions negatively affecting the decision-making, since the evaluation of structural improvement[23] is assessed against a 'frozen' target. Put differently, a member state will not be obliged to achieve even more structural progress if the previous year's output gap shifts upwards; however, a judgement on "exceptional circumstances" may take a revision in favour of the member state into account. Additionally, the so-called top-down approach of the excessive deficit procedure for evaluating the 'fiscal effort' of a member state adjusts the structural improvement for forecast errors, which then cannot be held against the member state. Lastly the European Commission has recently emphasised its commitment to make full use of the flexibility in the SGP – as also reflected in the recommendation to refrain from sanctioning Spain and Portugal in the summer of 2016. This critical case-by-case assessment lessens the focus on the mere numerical targets, which is prone to errors. Overall the SGP seem to be well equipped to deal with estimation errors to avoid costly ill-fitting policy advice and unwarranted sanctions.

It should be noted, however, that the flexibility and safeguards against revisions in the CAB are often asymmetric and only prevent strict application in cases of violations and unwarranted sanctioning, whereas (ex post) neglected sanctions or deserved ex post recommendations to tighten fiscal policy are not accounted for in the same manner. This could lead a country to accumulating debt over the cycle despite having fulfilled the official requirements (let alone making use of the additional flexibility). In order to move closer to the aim of debt sustainability, it might thus be useful to create an additional buffer. Our data suggest a general downward bias in the CAB (i.e. larger deficits). One simple solution would be to change the downward bias of rounding to the nearest 0.25% GDP of the MTO to rounding in the opposite direction. Another option would be to adjust the MTO target for the average revisions observed for that country over the past 10 years. The latter will be politically difficult and perhaps not proportionate to the problem at hand.

Even without these proposed changes, however, the SGP is faring quite well in dealing with CAB revisions. Its performance in this regard should be continuously monitored over the next

---

[23] In the form of cyclically-adjusted debt reduction.

decade since the analysis would greatly benefit from a longer time series that would encompass additional complete economic cycles.

Sources

Alcidi, C., D. Gros and G. Thirion (2016), "Is fiscal policy pro-cyclical in bad times?", FIRSTRUN project working paper, available at: http://www.firstrun.eu/files/2016/03/D4.1_fiscal_policy.pdf [Accessed: 17.06.2016].

Beetsma, R., B. Bluhm, M. Giuliodori and P. Wierts (2013), "From Budgetary Forecasts to Ex Post Fiscal Data: Exploring the Evolution of Fiscal Forecast Errors in the European Union", *Contemporary Economic Policy*, Volume 31(4), pp. 795–813, October

Beetsma, R., M. Giuliodori and P. Wierts (2009), "Planning to Cheat: EU Fiscal Policy in Real Time", *Economic Policy*, 24(60), pp. 753–804.

Beetsma, R., M. Giuliodori, M. Walschot, and P. Wierts (2010), "Fifty Years of Fiscal Planning and Implementation in the Netherlands", CEPR Discussion Paper No.7969, Centre for Economic Policy Research, London.

Benetrix, A. and P. Lane (2013), "Fiscal cyclicality and EMU", *Journal of International Money and Finance* 34, pp. 164–176.

Bernoth, K. and A. Hughes-Hallett and J. Lewis (2015), "The Cyclicality of automatic and discretionary fiscal policy: What can Real-time data tell us?", *Macroeconomic Dynamics*, Vol. 19, pp. 221–243.

Bernoth, K., A. Hughes-Hallett, and J. Lewis (2008), "Did Fiscal Policymakers Know What They Were Doing? Reassessing Fiscal Policy with Real-Time Data", CEPR Discussion Paper No. 6758, Centre for Economic Policy Research, London.

Bouis, R., B. Cournede and A. Christense, (2012), "Implications of Output Gap Uncertainty in Times of Crisis", OECD Economics Department Working Papers No. 977, OECD Publishing, Paris.

Buti, M. and P. van den Noord (2004), "Fiscal policy in EMU: Rules, discretion and political incentives," European Commission, *European Economy* - Economic Papers 206.

Brück, T. and A. Stephan (2006), "Do Eurozone Countries Cheat with their Budget Deficit Forecasts?", DIW Discussion Paper 508, German Institute for Economic Research, Berlin, 24 May.

Cebotari, A, J. M. Davis, L. Lusinyan, A. Mati, P. Mauro, M. Petrie, and R. Velloso, (2009), "Fiscal Risks. Sources, Disclosure, and Management", Fiscal Affairs Department, IMF, Washington, D.C.

Cimadomo, J. (2007), "Fiscal Policy in Real Time", CEPII Working Paper No. 07-10, Centre d'Études Prospectives et d'Informations Internationales, Paris.

Cimadomo, J. (2011), "Real-Time Data and Fiscal Policy Analysis: A Survey of the Literature", ECB Working Paper No. 1408, December.

Cimadomo, J. (2012), "Fiscal Policy in Real Time", Scand. J. of Economics, Vol. 114(2), pp. 440–465.

Croushore, D. (2011), "Frontiers of Real-Time Data Analysis", *Journal of Economic Literature*, Vol. 49, pp. 72–100.

Croushore, D. and S. van Norden (2014), "Fiscal Policy: Ex Ante and Ex Post", Federal Reserve Bank Philadelphia Working Paper No. 14-22, Philadelphia, PA, August.

Croushore, D., and T. Stark (2001), "A Real-Time Data Set for Macroeconomists", *Journal of Econometrics*, Vol. 105(1), pp. 111–130.

D'Elia, A. (2015), "EU Fiscal Framework", Presentation at the 2015 CEPS Summer School,

8 September 2015, Brussels.

De Castro, F., J. J. Perez and M. Rodriguez-Vives, (2011), "Fiscal Data Revisions in Europe", ECB Working Paper, No. 1342, European Central Bank, Frankfurt.

Egert, B. (2010), "Fiscal Policy Reaction to the Cycle in the OECD - Pro- or Counter-cyclical?", OECD Economics Department Working Papers, No. 763, OECD, Paris.

European Commission (2012a), "Consolidated Version of the Treaty on the Functioning of the European Union", *Official Journal of the European Union* C 326/49.

European Commission (2012b), "Specifications on the implementation of the Stability and Growth Pact and Guidelines on the format and content of Stability and Convergence Programmes", 3 September.

European Commission (2015), "Making the best use of the flexibility within the existing rules of the Stability and Growth Pact, COM(2015) 12 final, 13 January.

European Commission (2016a), "Vade Mecum on the Stability and Growth Pact", Institutional paper 021, March.

European Commission (2016b), "Making it happen: the European Semester", available at: http://ec.europa.eu/europe2020/making-it-happen/index_en.htm [Accessed: 16.01.2016].

European Commission (2016c), "Stability and Growth Pact: fiscal proposals for Spain and Portugal", EC Press Release, Brussels, 27 July 2016.

European Commission (2016d), "Recommendation for a Council Decision giving notice to Spain to take measures for the deficit reduction judged necessary in order to remedy the situation of excessive deficit", SWD(2016) 263 final, 27.7.2016.

European Commission (2016e), "Recommendation for a Council Decision giving notice to Portugal to take measures for the deficit reduction judged necessary in order to remedy the situation of excessive deficit", SWD(2016) 264 final, 27.7.2016.

European Commission (2016f), "AMECO Database", available at: http://ec.europa.eu/economy_finance/db_indicators/ameco/zipped_en.htm, [Accessed: 06.06.2016].

European Commission (2016g), "Statistical Annex of European Economy" http://ec.europa.eu/economy_finance/db_indicators/statistical_annex/index_en.htm, [Accessed: 06.06.2016].

European Council (2011), "Council Regulation (EU) No 1177/2011 of 8 November 2011 amending Regulation (EC) No 1467/97 on speeding up and clarifying the implementation of the excessive deficit procedure", Official Journal of the European Union, L 306/33, 23.11.2011.

European Council (2012), "Treaty on Stability, Coordination and Governance", available at: http://www.consilium.europa.eu/european-council/pdf/Treaty-on-Stability-Coordination-and-Governance-TSCG/, [Accessed: 11.02.2016].

Forni, L. and S. Momigliano (2005), "Cyclical Sensitivity of Fiscal Policies Based on Real-Time Data", *Applied Economics Quarterly*, 50(3), pp. 299–326.

Frankel, J. (2011), "Over-optimism in Forecasts by Official Budget Agencies and Its Implications", Review of World Economics, Vol. 149 (2), June 2013, pp. 247–272.

Frankel, J. and J. Schreger (2013), "Over-optimistic official forecasts and fiscal rules in the eurozone", NBER Working Paper No. 18283, National Bureau of Economic Research, Cambridge, MA.

Golinelli, R. and S. Momigliano (2006), "Real-Time Determinants of Fiscal Policies in the Euro Area", *Journal of Policy Modeling*, Vol. 28(9), pp. 943–964.

Gordo-Mora, L. and J. Nogueira-Martins (2007), "How reliable are the statistics for the Stability and Growth Pact?", *European Economy* - Economic Papers 2008-2015 273, Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Havik, K., K. Mc Morrow, F. Orlandi, C. Planas, R. Raciborski, W. Röger, A. Rossi, A. Thum-Thysen and V. Vandermeulen (2014), "The Production Function Methodology for Calculating Potential Growth Rates & Output Gaps", Economic Papers 535, November.

Holm-Hadulla, F., S. Hauptmeier and P. Rother (2010), "The Impact of Expenditure Rules on Budgetary Discipline over the Cycle", ECB Working Paper Series, No. 1169, April.

Hughes-Hallett, A., R. Kattai and J. Lewis (2011), "How reliable are cyclically adjusted budget balances in real time?", *Contemporary Economic Policy*, Vol. 30(1), pp. 75–92.

Jong-A-Pin, R., J-E Sturm and J. de Haan (2012), "Using Real-Time Data to Test for Political Budget Cycles", CESifo Working Paper Series 3939, CESifo Group, Munich.

Jonung, L. and M. Larch (2006), "Improving fiscal policy in the EU: The case for independent forecasts", Economic Policy CEPR, CES, MSH, Vol. 21(47), pp. 491-534.

Kempkes, G. (2014), "Cyclical Adjustment in Fiscal Rules: Some Evidence on Real-Time Bias for EU-15 Countries", *Public Finance Analysis*, Vol. 70(2), pp 278-315.

Larch, M. and A. Turrini (2010), "The Cyclically Adjusted Budget Balance in EU Fiscal Policymaking", *Intereconomics* 45(1), pp. 48–60.

Ley, E. and F. Misch (2014), "Real-Time Macro Monitoring and Fiscal Policy", ZEW Discussion Paper No. 14-122, Zentrum für Europäische Wirtschaftsforschung/Center for European Economic Research, Mannheim.

Marinheiro, C. (2011), "Fiscal sustainability and the accuracy of macroeconomic forecasts: Do supranational forecasts rather than government forecasts make a difference?", *International Journal of Sustainable Economy* 3(2), pp. 185-209, January.

Massimiliano, M. and A. Musso (2011), "The Reliability of Real-time Estimates of the Euro Area Output Gap", *Economic Modelling,* Vol. 28, pp. 1842–1856.

Merola, R. and J. Perez (2013), "Fiscal forecast errors: Governments versus independent agencies?", *European Journal of Political Economy,* Vol. 32, pp. 285–299.

Moulin, L. and P. Wierts (2006), "How Credible are Multiannual Budgetary Plans in the EU?", *Fiscal Indicators*, Banca d'Italia.

Mourre, G., G. M. Isbasoiu, D. Paternoster and M. Salto (2013), "The cyclically-adjusted budget balance used in the EU fiscal framework: an update", Economic Papers 478, March.

Mourre, G., C. Astarita, S. Princen (2014), "Adjusting the budget balance for the business cycle: the EU methodology", Economic Papers 536, November.

Odor, V. and J. Kucserová (2014a), "Finding Yeti - More robust estimates of output gap in Slovakia", CBR Working Paper No. 2/2014, National Bank of Slovakia, Bratislava.

Orphanides, A. (2001), "Monetary Policy Rules Based on Real-Time Data", *American Economic Review*, Vol. 91(4), pp. 964–985.

Orphanides, A., and S. van Norden (2002), "The Unreliability of Output Gap Estimates in Real Time", *Review of Economics and Statistics*, Vol. 84, pp. 569–583.

Paloviita, M. and P. Ikonen (2016), "How to explain errors in budget balance forecasts in euro area countries? Real-time evidence based on IMF data", Bank of Finland Research Discussion Paper No. 17/2016, Helsinki.

Pina, A. (2009), "Elusive Counter-Cyclicality and Deliberate Opportunism? Fiscal Policy from Plans to Final Outcomes", Bank of Portugal Working Paper, No. 6, Lisbon, April.

Pina, A. and N. Venes (2011), "The Political Economy of EDP Fiscal Forecasts: An Empirical Assessment", *European Journal of Political Economy*, Vol. 27(3), pp. 534–546, September.

Sorensen, P.B. and H. J. Whitta-Jacobsen (2010), "Introducing Advanced Macroeconomics: Growth and Business Cycles", Second Edition, McGraw-Hill Education, Columbus.

Tereanu, E., A. Tuladhar and A. Simone (2014), "Structural Balance Targeting and Output Gap Uncertainty", IMF WP/14/107, International Monetary Fund, Washington, D.C..

Virkola, T. (2014), "Real-Time Measures of the Output Gap and Fiscal Policy Stance", ETLA Report 37, Research Institute of the Finnish Economy, Helsinki.

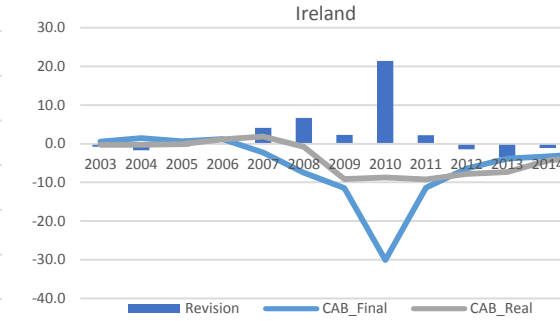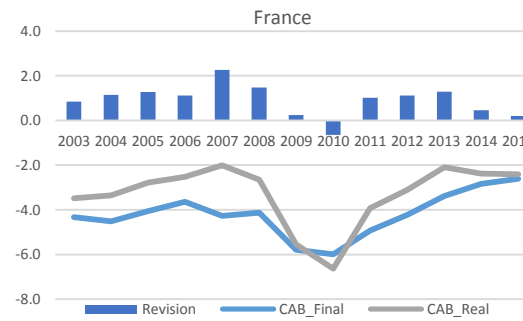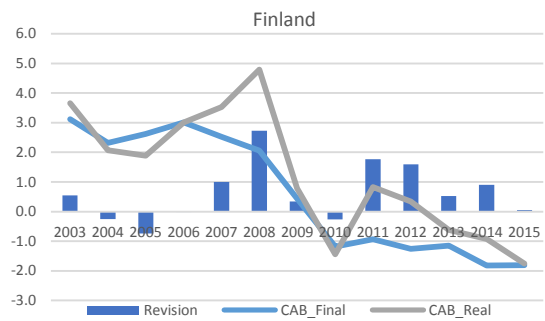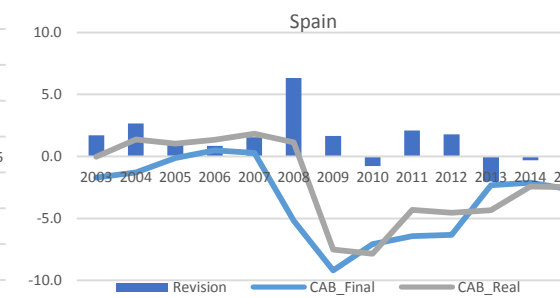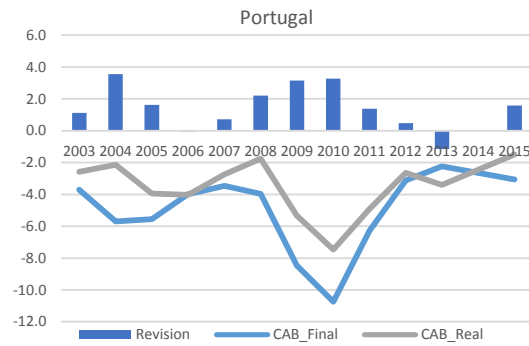Von Hagen, J. (2010), "Sticking to Fiscal Plans: The Role of Fiscal Institutions", *Public Choice*, 144, pp. 487–503.
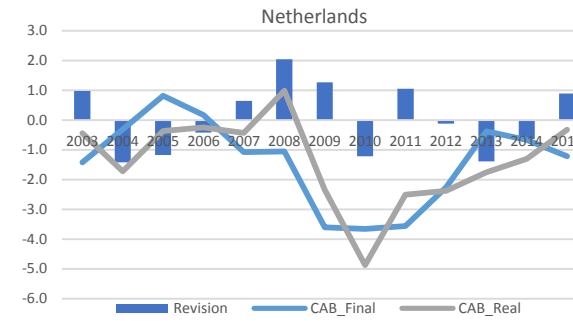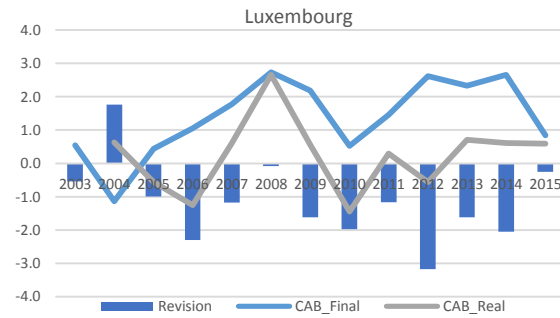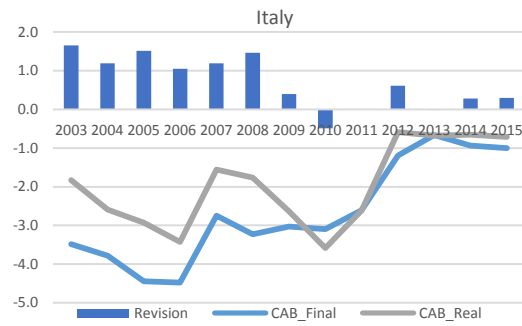
Appendix

Appendix 1. The European Semester
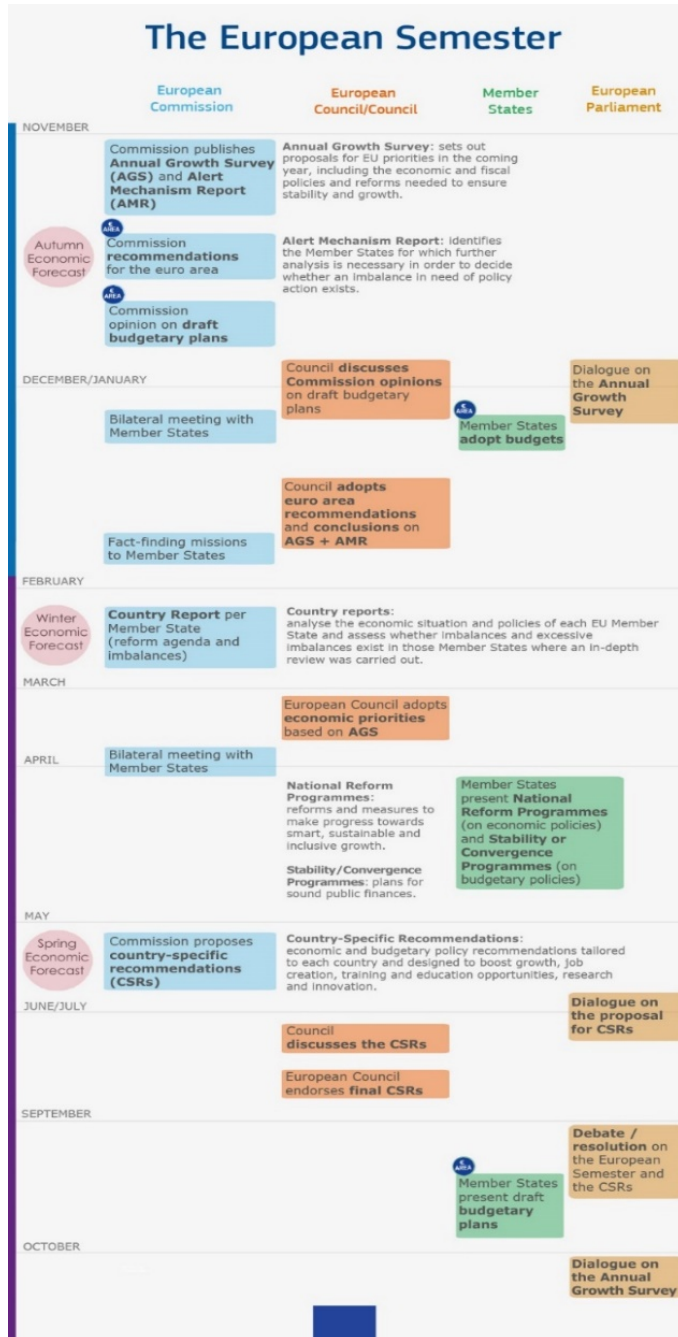
Source: European Commission (2016b)

Appendix 2. Country overviews
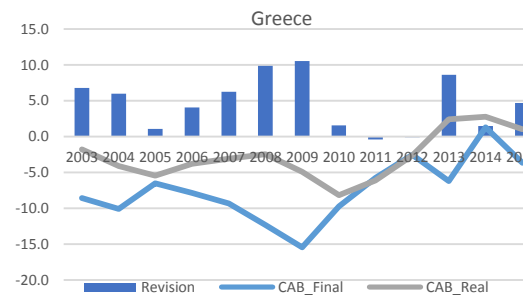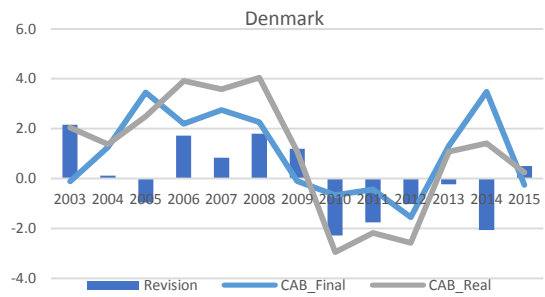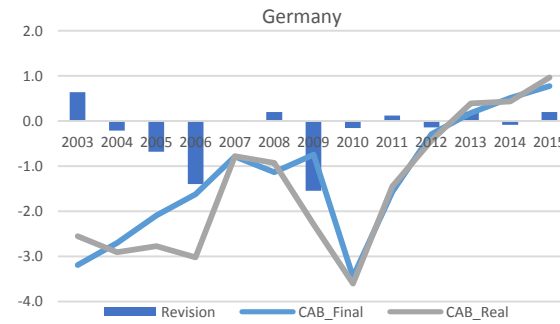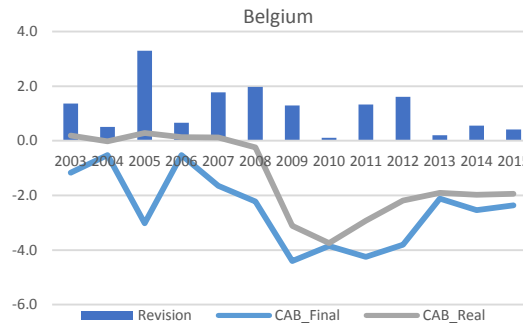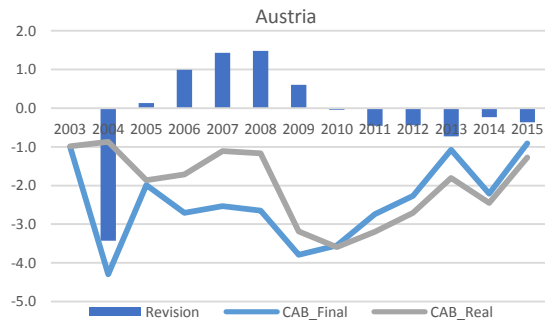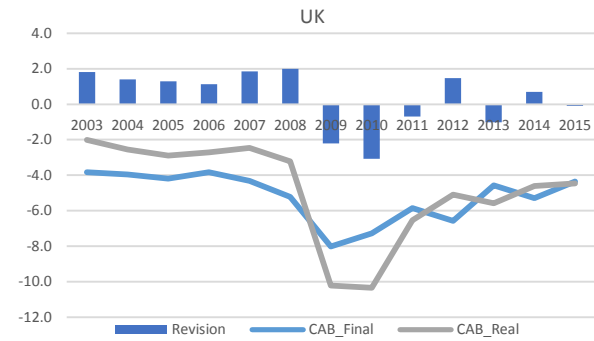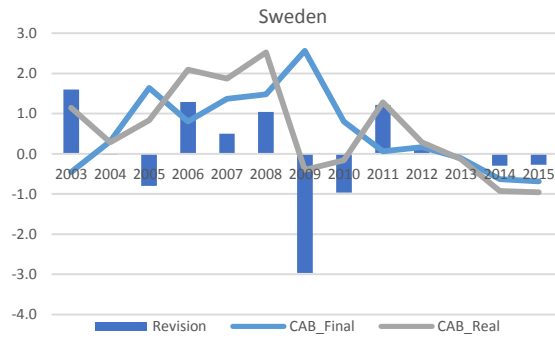
Appendix 3: Dynamics of CAB revisions

$y = 0.545x - 0.402$
$R^2 = 0.417$
CAB t-1 spring / CAB final

$y = 0.587x - 0.178$
$R^2 = 0.528$
CAB t-1 autumn / CAB final

$y = 0.681x - 0.061$
$R^2 = 0.641$
CAB t spring / CAB final

$y = 0.806x + 0.107$
$R^2 = 0.788$
CAB t autumn / CAB final

$y = 0.940x + 0.453$
$R^2 = 0.866$
CAB t+1 spring / CAB final

$y = 0.974x + 0.476$
$R^2 = 0.913$
CAB t+1 autumn / CAB final

$y = 0.978x + 0.412$
$R^2 = 0.923$
CAB t+2 spring / CAB final
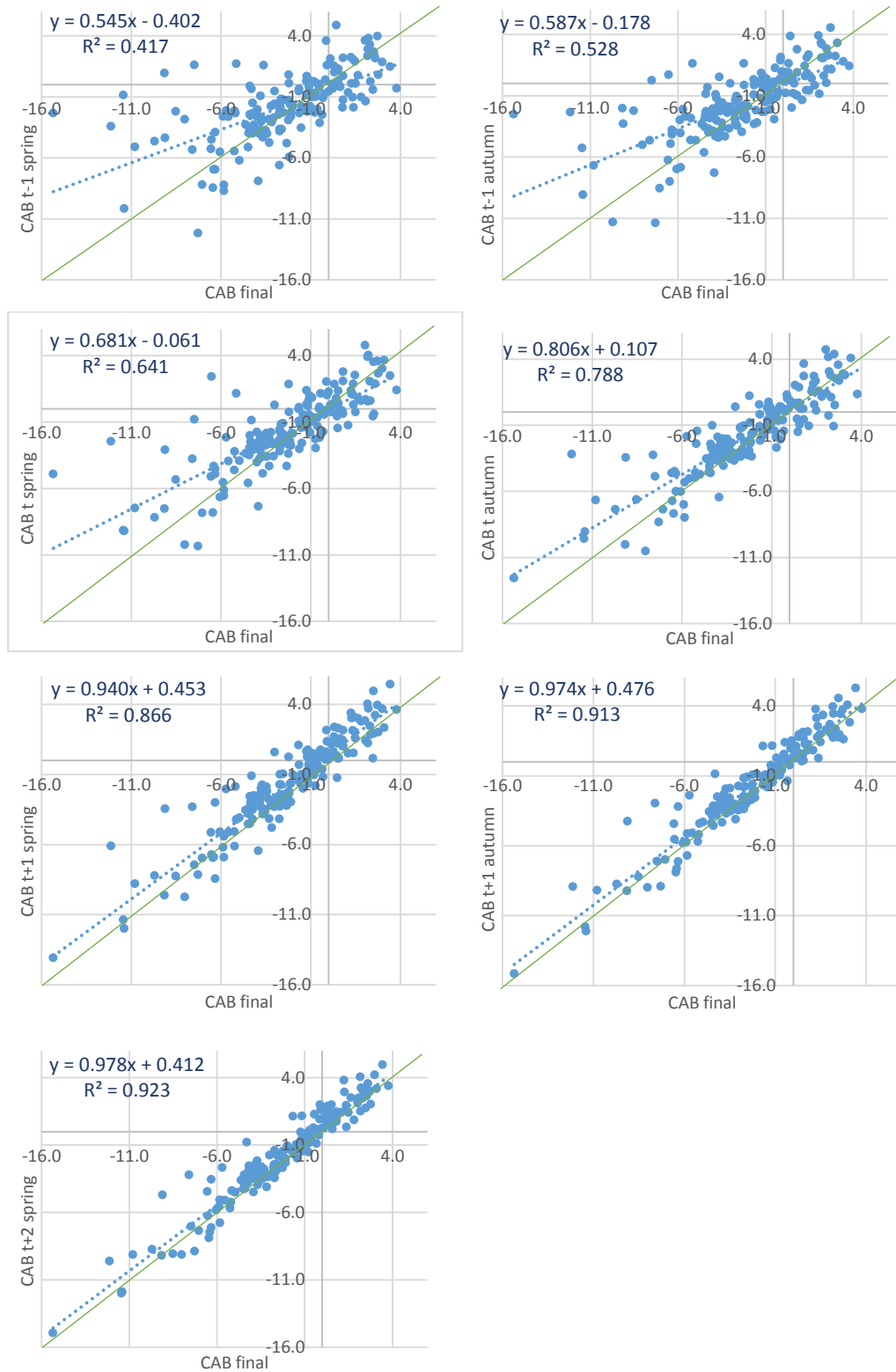
Appendix 4. Revisions of the CAB by year, 2003-13