WP/14/107

IMF Working Paper

Structural Balance Targeting and Output Gap Uncertainty

Eugen Tereanu, Anita Tuladhar, and Alejandro Simone

INTERNATIONAL MONETARY FUND

 WP/14/107

IMF Working Paper

European Department

Structural Balance Targeting and Output Gap Uncertainty

Prepared by Eugen Tereanu, Anita Tuladhar, and Alejandro Simone[1]

Authorized for distribution by James Daniel

June 2014

Abstract

Potential output estimation plays a crucial role in conducting fiscal policy based on structural balances. Difficulties in estimating potential output could lead to an erroneous policy stance with a consequent impact on growth. This paper analyzes historical data on revisions of actual and potential growth in the European Union and the implication of these revisions for the measurement of fiscal effort using the cyclically-adjusted primary balance (CAPB). It finds that revisions in output gap estimates were large, at almost 1½ percent of potential GDP on average. Revisions in potential GDP also contributed significantly to revisions in the estimated CAPB, especially during the crisis years. Given these findings and historical correlations, it proposes an indicative rule of thumb for reducing errors in the measurement of fiscal effort by factoring in that about 30 percent of revisions in actual growth capture changes in potential growth. In other words, the standard advice of "letting automatic stabilizers operate fully" in response to a positive/negative growth shocks likely implies a strengthening/weakening of the structural position.

JEL Classification Numbers: E62, H6
Keywords: structural balance, cyclically adjusted balance, fiscal policy

Authors' E-Mail Addresses: etereanu@imf.org; atuladhar@imf.org; asimone@imf.org;

[1] The authors are grateful to Michal Andrle, Helge Berger, James Daniel, Natan Epstein, Estelle Liu, Camelia Minoiu, Rodrigo Valdez, Daria Zakharova and seminar participants at the IMF for helpful comments and suggestions. We would also like to thank Geraldine Mahieu, Cristina Cheptea, and Sarika Chinta for assisting with the collection of the data, charts, and document formatting respectively.

Table of Contents **Pages**

I. Introduction

There has been an increasing interest in adopting a more structural approach to fiscal policy in Europe. In particular, the newly-approved EU Fiscal Compact requires countries to include structural targets as part of their fiscal rules, against which policy performance will be evaluated. Successful implementation of structural targets relies on the assumption that both actual and potential growth can be accurately estimated. In practice, however, this assumption rarely holds. Real-time estimates of both actual and potential growth are often subject to significant revisions. Such revisions may continue to take place beyond the horizon of the actual budget year. In addition, the unobservable potential growth is computed using different techniques to extract information about underlying economic trends. These methods have been evolving over time as illustrated by recent changes to the EC methodology for computing structural unemployment. Potential growth estimates may also not be straightforward to update on the basis of short-term economic developments.

This paper analyzes historical data on the revisions of actual and potential growth in the EU and focuses specifically on the following questions:

- How large were the errors in measuring the output gap?
- What were the implications of revisions in potential GDP for the measurement of fiscal effort?
- Can an indicative rule of thumb, based on the historical correlation, be observed between revisions in actual and potential growth?

Our results suggest the following answers:

- ***Output gap errors (as measured by the revisions) were large.*** Although the cyclical position suggested by the sign of the output gap has changed infrequently as a result of output gap revisions, the *magnitudes* of output gap revisions have been significant on average. They averaged about 1½ percent of potential GDP in absolute terms between the time when budget assumptions were made and when the outturn data was first evaluated. Revisions during non-crisis years were also sizable, averaging almost 1 percent of potential GDP.

- ***Fiscal effort estimates as measured by the cyclically-adjusted primary balance (CAPB) were also revised substantially.*** Potential GDP contributed significantly to these revisions, especially during crisis years when it was the largest contributing factor (e.g., more than the deviation from target of key fiscal variables such as primary expenditure). As a result, estimates of fiscal effort based on the CAPB have been subject to considerable variation and error due to output gap and potential output revisions.

- ***On average about 30 percent of the revision in actual growth reflected revisions in potential growth.*** A simple rule of thumb which takes into account the historical correlation between revisions in actual output growth and potential output growth could help reduce errors in the measures of fiscal effort using CAPB due to output gap and potential output revisions. These finding suggests that short-run growth revisions should not be viewed as solely cyclical but rather also including a permanent structural component due to changes in potential output. In other words, the standard advice of "letting automatic stabilizers operate fully" in response to a positive/negative growth shock likely implies a strengthening/weakening of the structural position.

The paper is organized as follows. Section II presents a literature review. Section III discusses the impact of output gap revisions on fiscal effort. Section IV analyses the correlation between growth surprises and potential growth. Section V concludes.

II. Literature Review

The uncertainty in accurately measuring the output gap is well documented in the literature. In the context of monetary policy rules, Orphanides and van Norden (2002, 2005) documented the challenges of using the output gap as a measure of the economic slack in the policy reaction function. Using US data from 1969 to 2003, they show that revisions in the real-time estimates of output gap are as large as the historical estimates themselves. They construct output gaps under alternative methodologies and find that for many of these methods, the signs even differ between historical and real-time estimates. Using a similar methodology, Cayen and van Norden (2005) find that revisions in Canadian output gaps are also significant. In contrast to the US data, where the main source of revisions is the unreliability of end-of sample estimates of the trend in output, data revision was the main reason for the change in the Canadian output gap data. They further show that using the change rather than the level of the output gap reduces the measurement problem only modestly.

Several recent papers have examined the uncertainty in calculating output gaps in Europe. Using data since 1999 for euro area countries, Massimiliano and Musso (2011) show that ex-post revisions of the estimated output gaps are of the same order of magnitude as the estimates of output gaps themselves. The uncertainty is mostly attributable to parameter instability, model uncertainty, and unreliable end-of-sample estimates of the trends in output. Data revisions play a relatively minor role. The uncertainties are particularly acute during turning points where accurate estimates would be particularly useful for policy-making.

The literature also points to non-trivial errors in measuring CAPBs. Koske and Pain (2008) show that over the period 1995–2003, on average across the G7 countries and also for 21 OECD economies, output and unemployment gap revisions accounted for revisions of around 0.4 percentage points of GDP in the CAPB. Ex-post revisions for output gap estimates are larger particularly around turning points. In a sample of OECD countries, Bouis and others (2012) find that revisions of 1 to 1.5 percentage points to output gaps are relatively common although underlying fiscal balances are reasonably robust to such revisions. On the other hand, Hallett and others (2009) find that revisions of CAPBs are over 1 percent of GDP in most euro area countries owing to revisions in estimates of output gap. They note that real-time CAPBs have low power in detecting fiscal slippages and in correctly identifying fiscal improvements. Moreover, real time CAPBs are systematically less reliable under conditions of poor or deteriorating public finances. Jonung and Larch (2006) find that due to a systematic upward bias in government produced forecasts of potential output, real time assessments of fiscal position have been overly optimistic. Using WEO data for 175 countries, Ley and Misch (2013) also find that revisions in output have substantial effects on the estimation of structural balance, which can imply substantial debt accumulation.

This paper contributes by assessing the implications of measurement error on the cyclically adjusted primary balances using recent data for the EU countries. It also compares ex-post data with real time data available at the time of budget preparation and implementation to assess fiscal policy implications generated by these measurement problems. As an operational tool to implement structural fiscal policy, it also proposes a rule of thumb on the relation between output surprises and potential growth based on this historical correlation.

III. Evaluating Fiscal Effort: The Impact of Output Gap Revisions

Changes in CAPB, often used as a proxy for assessing fiscal effort, are subject to measurement bias. In calculating CAPB, the underlying fiscal position is estimated by stripping away the impact of the economic cycle from fiscal variables. Therefore, CAPB can change not only due to variation in fiscal variables but also due to changes in the output gap arising from revisions in potential output. In this section, we examine the extent to which CAPB estimates have been revised between the time of budget planning and ex-post after the budget has been executed. We also assess the impact of revisions in potential output on CAPB estimates by decomposing the total change in the CAPB due to revisions in its different components and a residual term. We first describe the data and subsequently the methodology for the decomposition.

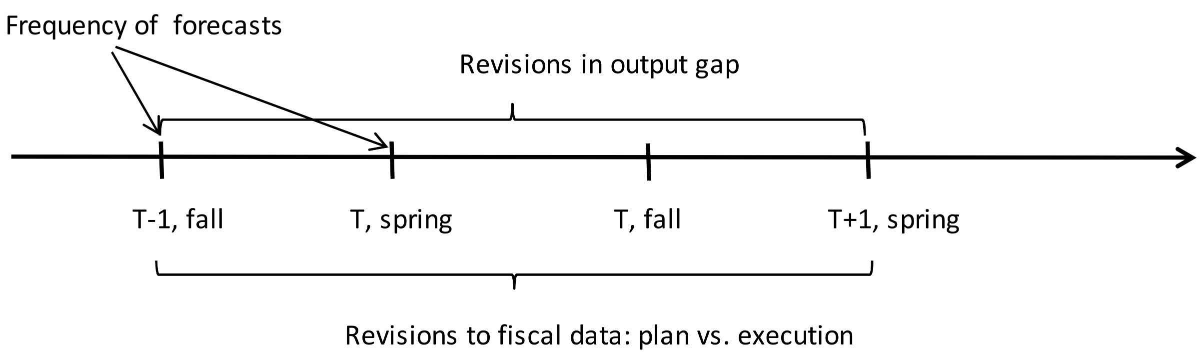
A. Data and Methodology

Data

The data for the analysis is derived from the European Commission's semi-annual (fall and spring) economic forecasts covering the period 2003–12. The fact that this dataset is used for policy making purposes in the context of the EU fiscal compact makes its analysis particularly relevant. Moreover, the advantage of this dataset is that, unlike others, it has been computed consistently across countries using a uniform estimation methodology for potential output[2]. This allows to reduce potential sources of variation in the data and focus on the variation in the forecast and outturn data for output, potential output and the output gap across different forecast vintages. Data for general government revenues, expenditures and GDP deflator are also obtained from the EC's Economic Forecasts. The analysis was also done using World Economic Outlook data for the EU countries. Since methodologies for estimating potential output vary across time and countries, results are not presented in this paper. We find that revisions in potential output and output gap are also sizable in this other data source.

Time horizons for defining revisions

Figure 1. Time horizon for evaluating fiscal stance for budget year *t (budget horizon)*



The budget evaluation horizon. To analyze the implications of output gap revisions on the measured fiscal effort, we define a *budget evaluation horizon,* over which we estimate the revision in both output gap and the CAPB, as the difference between the estimate at the budget preparation stage (t-1,f) (the forecast vintage in the fall preceding budget year *t*) and the first budget evaluation stage (t+1,s) (the forecast vintage in the spring of the year following budget year *t*). This is the difference between the first estimate for the budget year *t* produced during the spring of the ***following*** *year (t+1)* (when actual GDP and budget execution data for year *t* are first available) and the initial budget preparation forecast produced in the fall *of the* ***preceding*** *year (t-1).*

Methodology for decomposing the revision to the cyclically adjusted primary balance

We decompose the *cross-vintage change* in the CAPB ratio to GDP over the budget evaluation horizon into the contributions of individual sub-components (i.e. nominal

[2] The European Commission produces potential output series based on both the Hodrick-Prescott filter and the Production Function method. We use the latter estimates, which are underpinning the estimates of the cyclically adjusted balances presented in the general government data forecast tables.

government spending, nominal interest spending, GDP deflator, output and potential output) and a residual cross-term. The decomposition is done by allowing one of the variables to change while keeping all the others constant. For example, in the case of real potential GDP, $Y^{*}_{v,t}$, we assess its relative contribution to revisions of the CAPB by taking the difference between the initial CAPB and a new CAPB calculated using the revised potential GDP but keeping the initial values for the other variables constant (see Appendix I). In a similar way, we also calculate the contribution of other subcomponents. The only difference is with respect to the contribution of output revisions since the impact of output on the CAPB to GDP ratio is through two channels: the denominator effect as well as indirectly though the cyclical effect on revenues. Therefore, to calculate the contribution of output, we take into account the endogenous impact on revenues.

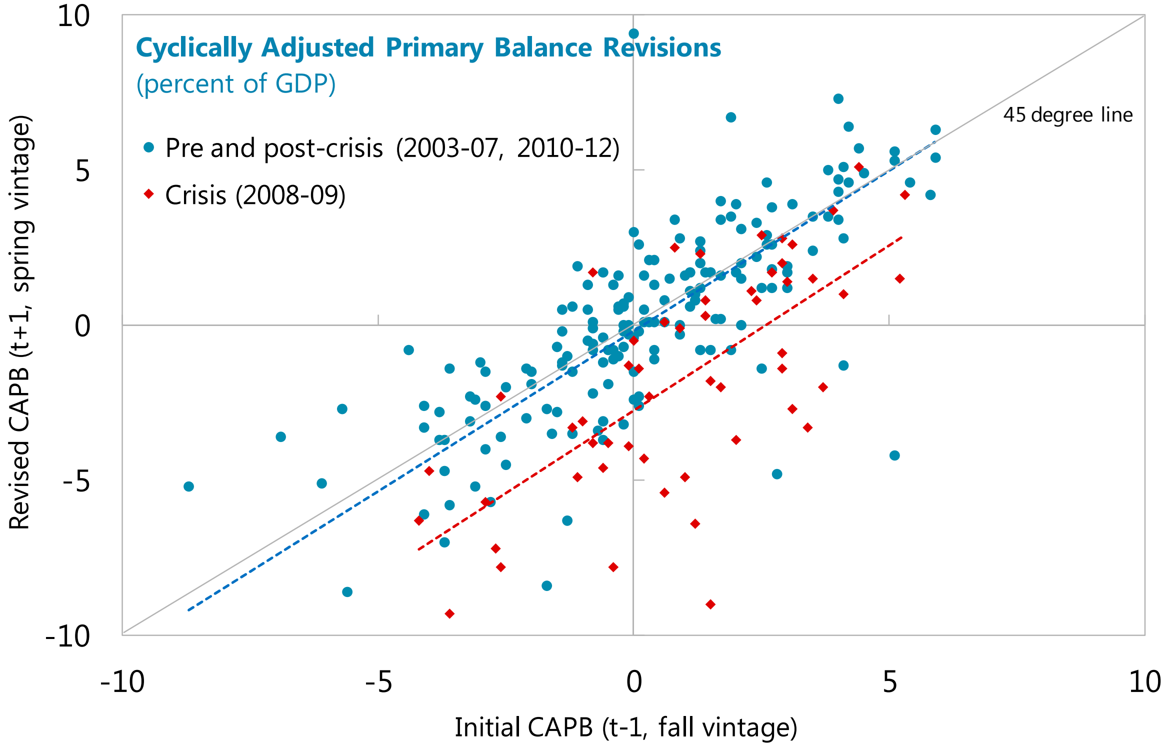
B. The Magnitude of Output Gap Revisions and Implications for the Measurement of CAPB

Revisions to CAPB

Estimates of CAPB have been revised substantially between planned budget and actual outcomes, particularly during crisis periods. This is shown by the distance between the observations and the 45 degree line in Figure 2. The x-axis shows the initial projected CAPB at the time of budget preparation and the y-axis shows the revised CAPB after the budgetary execution. Observations on the 45 degree line suggest the projected CAPB is in line with the ex-post outturn data. When we divide the sample into non-crisis periods (blue) and crisis periods (red), we see that the fitted line through the non-crisis observations is close to the 45 degree line, while the fitted line through the crisis-observations is a downward shift from the 45 degree line. In other words, during crisis years, the estimates of CAPB were considerably worse for most countries after they were re-estimated following budget execution compared to the budget preparation period. During non-crisis periods, revisions of CAPB estimates across countries varied, such that the average across countries was broadly similar to initial estimate of the CAPB.

In terms of magnitude, the size of these revisions was large during both crisis and non-crisis periods. The size of the mean absolute revision in CAPB over the budget horizon was almost 1¾ percent of GDP. During the peak crisis years of 2008 and 2009, mean absolute revision was about 3 percent of GDP. By comparison, in 2005–07 and 2010–11 (non-crisis years), the mean absolute revision was 1½ percent of GDP. The median value across countries for the mean absolute revision during the crisis period was 2.3 percent of GDP and during non-crisis period was 1 percent of GDP. Overall, the median value across countries for the absolute revision was 1¼ percent of GDP, suggesting that these revisions were not driven particularly by outliers.

Figure 2. During crisis years CAPBs were revised downwards for most countries



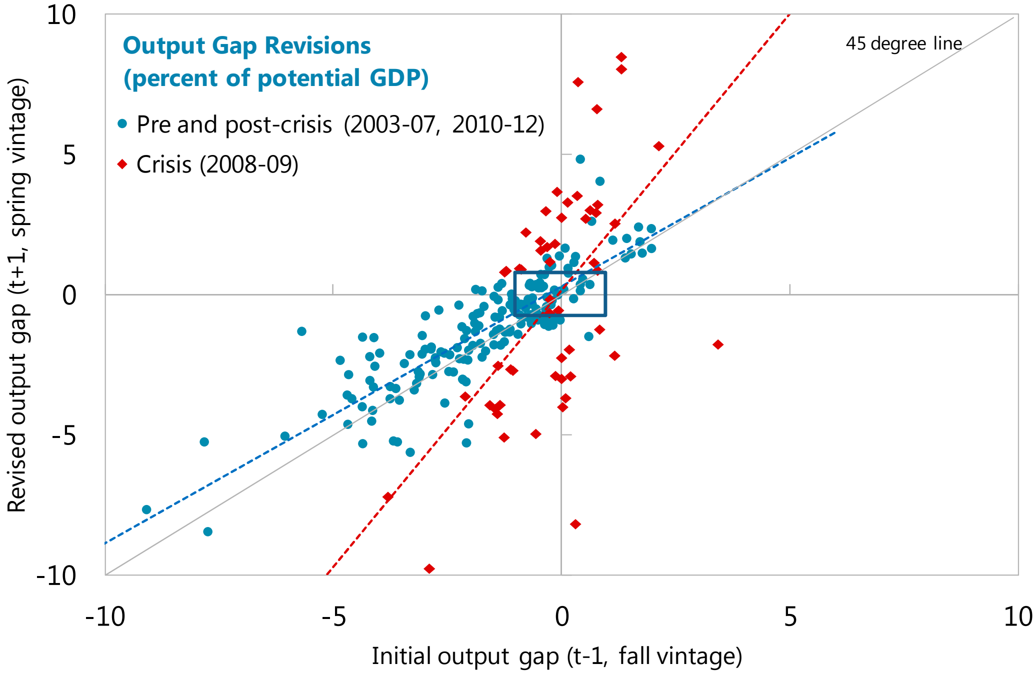
Sources: European Commission ; and IMF staff calculations.

These large revisions could reflect not just changes in the fiscal effort, but also measurement errors arising from revisions in potential output and output gap. We analyze below how large were the revisions in output gap and whether such revisions in output gap have contributed significantly to the revisions in CAPBs.

Output gap revisions

A comparison over the budget horizon shows that the average size of output gap revisions was large, in particular during crisis years. A positive/negative output gap means that actual is above/below potential GDP. As above, we examined the revisions during the crisis and non-crisis years between budget preparation and outturn, captured by the distance from the 45 degree line in Figure 3. During crisis periods (red dots), the output gap was revised up significantly in absolute terms as shown by the slope of the fitted line higher than the 45 degree line. Conceptually, this means that the size of the output gap at the time of the budget was considerably smaller than what was estimated with ex-post data irrespective of the sign of the output gap. During non-crisis periods, on average across countries, output gap was revised down only slightly between budget preparation and outturn, as shown by the slope of the fitted line which was slightly below the 45 degree line.

Figure 3. The size of the output gap grew during crisis years irrespective of sign



Sources: European Commission; and IMF staff estiimates.

The size of output gap revisions was significant. Over the period 2003–12, the mean absolute value of revisions of the output gap for the budget horizon was 1.3 percent of GDP. The largest revisions of some 3 percent of GDP took place during the crisis years of 2008 and 2009 and a smaller revision of 0.8 percent of GDP on average during non-crisis years[3]. The median value across countries for the absolute revision in output gap stood at 0.9 percent of GDP (2.5 percent of GDP during crisis periods and 0.6 percent of GDP during non-crisis periods).

These revisions were, however, not large enough to change the sign of the output gap. In other words, the error in estimating the *cyclical position* across the budget horizon was minimal (i.e., the output gap changed signs only infrequently following the revision). In about one fifth of the cases the sign of the output gap changed (upper left and bottom right

[3] The revisions above are focused on the budget horizon since this is most relevant operationally for fiscal policy-making. For completeness, we also examined the size of revisions over longer time horizons and the qualitative results above continue to hold. We find that the output gap was also revised significantly beyond the budget horizon as output and potential output continued to be revised for another two years. The mean absolute error from the budget preparation vintage to the final vintage three years later was 2.6 percent of GDP, compared with the 1.3 percent of GDP revision over the budget horizon as noted above. These revisions reflect not only output surprises but also significant revisions in potential GDP. For a more detailed discussion of revisions to output gap over time beyond the budget horizon and the contributing factors, see Appendix II.

quadrants, Figure 3). However, when excluding observations where the initial output gap was less than 1 percent in absolute terms, as shown by the observations inside the boxed area in the Figure 3, the number dropped to less than 7 percent of total observations.

The revisions in output gap reflect not only GDP forecast errors, but also significant revisions in potential output. This is seen by the large downward revision in potential output during 2008 and 2009 (Table 1), which reflects the large revisions that occurred in most countries following the Lehman crisis. But even in 2007 and 2010, the revisions to potential output were of a sizable magnitude.

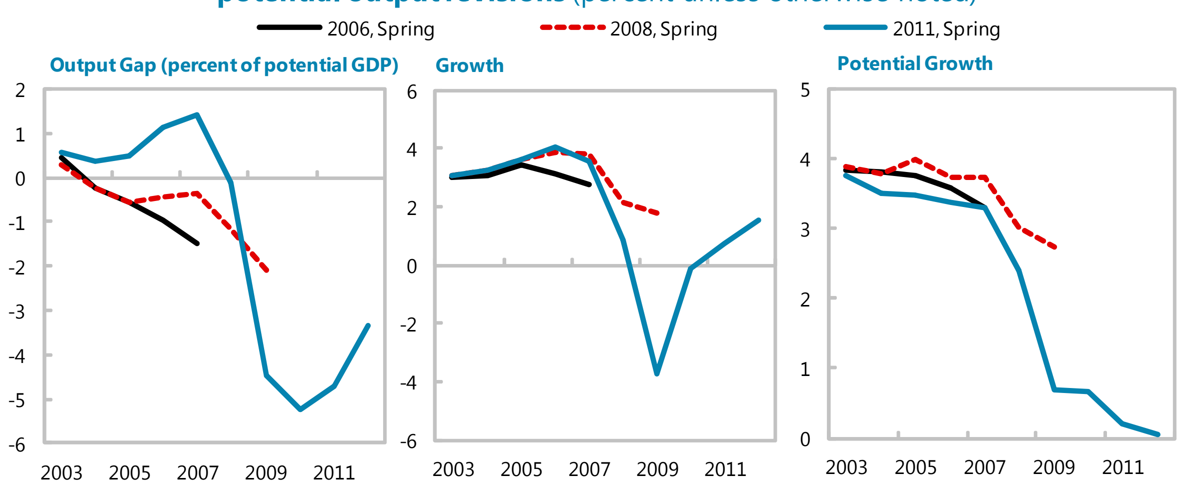
Table 1. EU: Revisions in Output and Potential Output over the Budget Horizon, Average

	Output Percent change	Potential Output Percent change	Output Gap Change in percentage points
2007	2.2	1.2	1.1
2008	-1.8	-4.5	2.9
2009	-7.3	-4.4	-3.0
2010	1.8	1.5	0.3

Sources: European Commission; and IMF staff estimates.

This can also be seen in the country example below which shows that the output gap was revised up successively over time. The first revision from 2006 to 2008 showed an upward revision to potential growth alongside a larger revision to GDP growth. The second revision between 2008 and 2011 shows a more positive output gap as potential growth is revised down, including for years going back to 2003.

Figure 4. Revisions in the Spanish output gap illustrate the large role played by potential output revisions (percent unless otherwise noted)



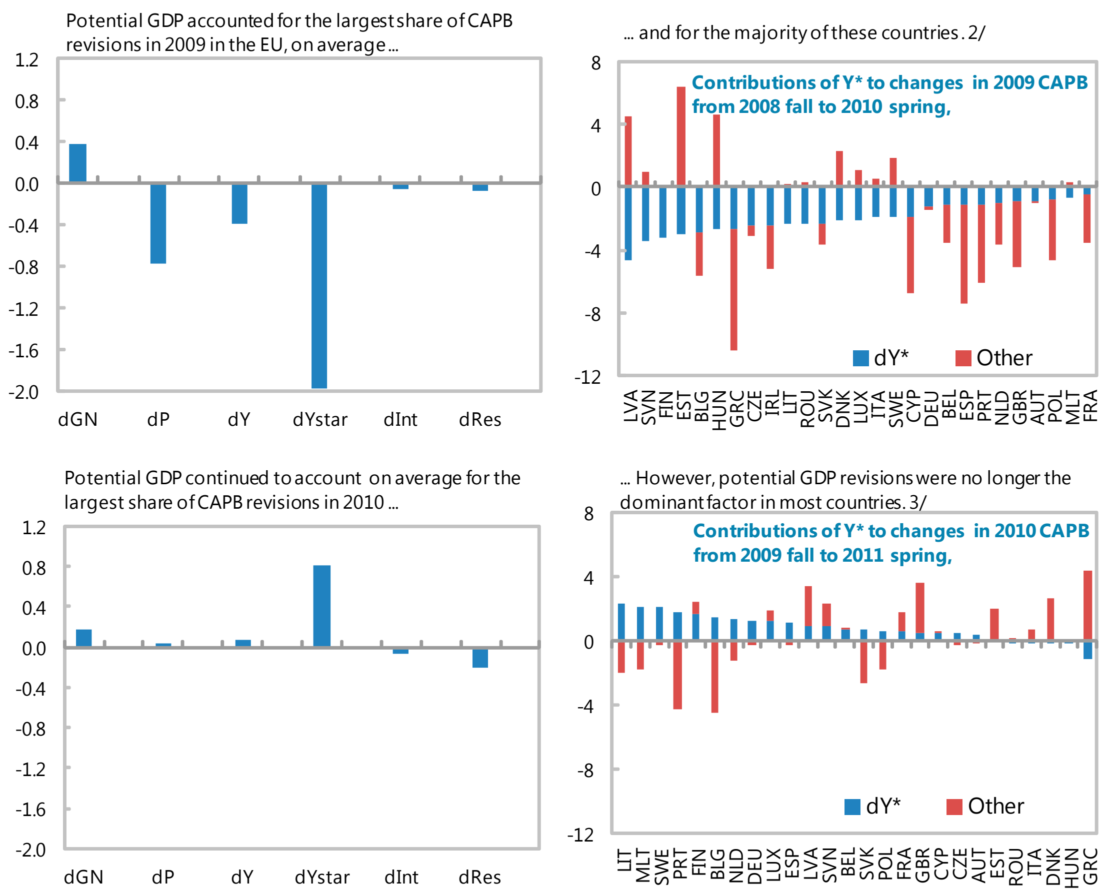
Source: European Commission

Decomposition of revisions to CAPB

Given the magnitude of output gap and potential GDP revisions, we decompose the change in the CAPB to determine the contribution of potential output to revisions in CAPB. The cross vintage change ("error") in CAPB is decomposed into contributions arising from forecast errors in output, potential output, GDP deflator, nominal level of primary government spending, interest payments and a residual term that captures the interaction between these variables. Since revenues are sensitive to the underlying output change, the contribution from the change in output includes the change in revenues as well.

The decomposition shows that revisions in estimates of potential GDP were a large contributing factor to the change in the CAPB (Figure 5). For each of the years 2009–10,

Figure 5. Revisions in estimates of potential GDP are a large factor contributing to revisions in the CAPBs (percent of GDP)



Source: European Commission and IMF Staff Estimates.

1/ Revisions defined for the following variabales :dGN = government primary expenditure; dP = GDP deflator; dY = output; dY* = potential output; dInt = interest payments; dRes = residual cross term.

2/ The sample of countries for 2009 includes the EU 27 countries: Latvia (LVA), Slovenia (SVN), Finland (FIN), Estonia (EST), Bulgaria (BLG), Hungary (HUN), Greece (GRC), Czech Republic CZE), Ireland (IRL), Lithuania (LIT), , Romania (ROM), Slovakia (SVK), Denmark (DNK), Luxembourg (LUX), Italy (ITA), Sweden (SWE), Cyprus (CYP), Germany (DEU), Belgium (BEL), Spain (ESP), Portugal (PRT), Netherlands (NLD), United Kingdom (GBR), Austria (AUT), Poland (POL), Malta (MLT), France (FRA).

3/ The sample of countries for 2010 includes the EU 27 countries excluding Ireland. The large revision of the CAPB for Ireland in 2010 was an outlier due to the inclusion of fiscal costs from the banking sector. To prevent distortion of the cross country averages, Ireland was excluded from the sample.

the revision in potential GDP is the largest contributor to the change in CAPB, more so than changes in primary spending. However, the aggregation masks large heterogeneity across countries. Cross country comparisons show that during 2009, the contribution of potential GDP is larger than other factors in majority of the cases. In 2010, revisions to potential were still significant but were not dominant in most of the cases. We repeated the decomposition for 2007 and 2008 and find similar results as for 2010 and 2009, respectively.

Overall, we find that the initial ex-ante "budget time" assessment on whether the economy was in a recession or above potential has not generally been changed ex-post. However, the large *size* of output gap revisions and the large role in them played by potential output revisions led to significant CAPB revisions. As a result, estimates of fiscal effort based on the CAPB have been subject to considerable uncertainty and error due to output gap and potential output revisions.

IV. Fiscal Policy Implications

The above results highlight the value of reliable estimates of potential output. In principle, approaches to estimate potential GDP that are less prone to forecast errors or rely on data that is less affected by revisions (e.g., surveys) offer a first-best solution to the problem of changing estimates of potential output and frequent recalculations of the CAPBs**.** However, given the inherent difficulties in accurately estimating an unobserved variable in real time and that most estimates of potential output—one way or the other—refer to GDP or its components, a second-best option for the practitioner may be trying to find a shortcut based on historical correlations between observed real time GDP data and final estimates of potential output growth. Specifically, a simple quantitative rule of thumb could provide a rough estimate of how much of a short term surprise in GDP growth rates will translate into changes in potential GDP. Such a rule of thumb would help reducing errors in estimating the output gap and potential output, and thus lead to reduced measurement errors of fiscal effort using CAPB estimates.

We estimate such a rule of thumb analyzing the historical correlation between short-term revisions in output growth and the long-term revision of potential growth (Appendix III)**.** We use a different horizon for this analysis because from a fiscal policy-making point of view, the main observed variable in real time is a short term surprise in GDP growth whereas the main variable of interest for calculating the "true" CAPB is the final estimate of potential growth. Therefore, to assess how much of the short-term, real-time growth revision could be interpreted as reflecting potential output revisions, the long-term revisions in potential growth (the measure for the final ,"true," potential growth revision) are regressed on a constant and the short-term, real-time growth revisions, using various country and time effects for both spring and fall growth revisions

We find that a 1 percentage point short-term output growth revision results on average in some 0.2–0.3 percentage point revision in potential growth in the long term in the same direction. This correlation is fairly robust across estimation methods using different country samples and time horizons and whether we use data from spring or fall revisions in the dataset. Further robustness checks on an out-of-sample basis are, unfortunately, not feasible given data limitations.

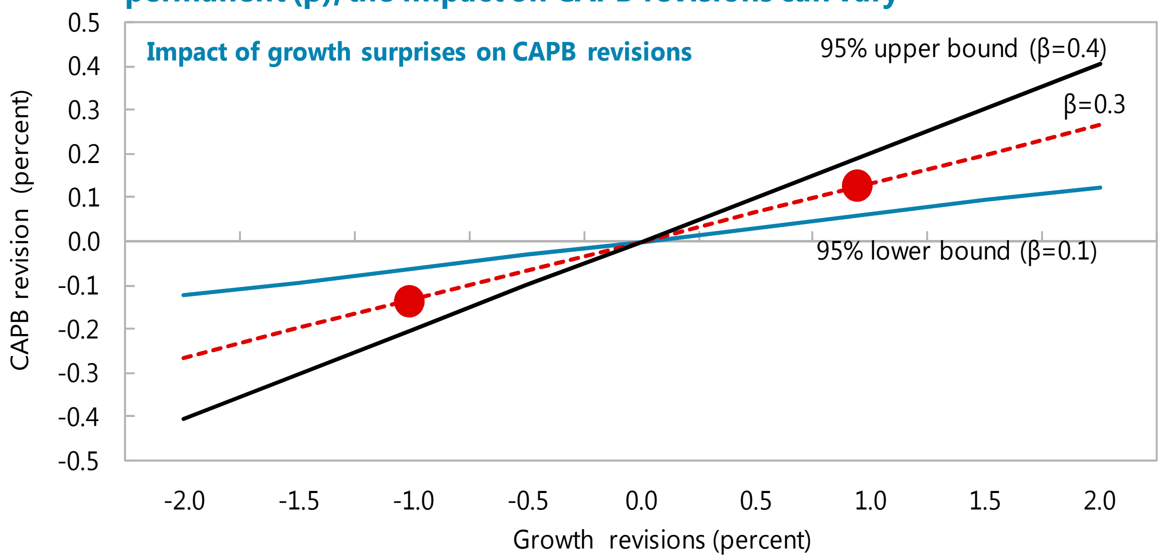
These findings suggest that short-run growth revisions include a revision of potential growth component that in turn impacts the estimation of the structural fiscal position. Thus, a positive/negative growth shocks imply a strengthening/weakening of the structural position. Our estimates suggest that, on average, about 30 percent of short-term growth revisions should be considered as reflecting a change in potential growth. This represents a rough rule of thumb on how much growth surprises should be considered as permanent when computing the output gap and therefore the CAPB. If fiscal policy objectives are based on structural balance targets, appropriate measures would be needed to compensate for the change in the structural balance given a certain structural balance objective.

To illustrate these findings, consider the following example (Figure 6):

- If growth is lower by one percent and budgetary semi-elasticity is 0.5, this would imply that the headline deficit would be higher by 0.5 percent of GDP. If this higher deficit reflected just cyclical effects, the structural balance would not change. However, since only 70 percent of the short-term growth revision is cyclical, one should expect the structural balance to have worsened by 30 percent or about 0.15 percent of GDP. Offsetting measures of the same magnitude would thus be needed to keep the structural balance constant, meaning that the headline deficit should be allowed to increase by only 0.35 percent of GDP in response to the short-term growth revision to keep the structural deficit unchanged.

- Symmetrically, if growth surprises on the upside, the headline deficit would fall by 0.5 percent of GDP. A structural improvement in the deficit of 0.15 percent of GDP

Figure 6. Depending on the share of a growth surprise that is permanent (β), the impact on CAPB revisions can vary



would thus be implied and permanent deficit increasing measures of this magnitude could be implemented without worsening the structural balance.

- In any given year, the impact of growth on potential growth and its consequent impact on CAPB could be higher or lower. The 30 percent rule of thumb described above (β=0.3) only represents the relationship on average over time. Depending on economic circumstances, judgment is required on how much of growth surprise reflects a supply shock and should be considered permanent. Thus, β may be larger or smaller than 0.3. For instance, during a crisis period, an economic recession could have a stronger permanent component and thus a larger β. The opposite may be true instead during normal periods. A 95 percentile range around the 0.3 average relationship suggests values for β between 0.1 and 0.4.

- In deciding the applicability of such a rule of thumb for a specific country, the CAPB estimate could also be compared with a "CAPB at Risk" measure. Such a measure would capture the country-specific structural balance estimate variance arising from forecast error variance in output and potential output.

These results can help strengthen structural balance rules. For example, several countries adopted rules that contain mechanisms to correct deviations from structural balance targets once they exceed a certain threshold, including to address potential output and output gap uncertainty. Taking into account the correlation between real time growth surprises and potential GDP growth will help assessing the nature of these deviations and calibrate the need for adjustment. The case for taking into account possible changes to potential output is particularly strong when growth surprises are large. In addition, because the rule of thumb takes into account the correlation between real-time actual growth surprises and long-term revisions to potential GDP growth, it can potentially improve the accuracy of real-time estimates of potential GDP growth.

V. Conclusion and Policy Implications

Successful implementation of structural fiscal targets requires the ability to measure potential output and output gaps accurately in real time. This is difficult since measuring an unobservable variable such as potential GDP is inherently subject to great uncertainty. While the first best approach would be to estimate potential GDP in a way that is less prone to forecast errors or relies on data that is less affected by revisions (e.g., surveys), an approximate rule of thumb could complement this approach by taking into account correlation between actual growth surprises and long-term revisions in potential growth. Using a historical database of the different vintages of estimates of output and potential output over the past decade for the EU, this paper suggests that:

- Potential GDP and the output gap are subject to significant revisions, especially during crisis periods. Such revisions have been a key factor in explaining revisions in CAPB estimates. Caution is therefore needed in interpreting CAPBs as an indicator of fiscal effort. These results highlight the need for structural balance rules that are robust to potential output and output revisions (i.e. those that include correction mechanisms for deviations from targets), while maintaining simplicity to ensure transparency.

- About a third of a growth surprise reflects on average a permanent change in potential growth. Such an indicative rule of thumb could help improve CAPB estimation and would imply that the standard advice of "letting automatic stabilizers operate fully" in response to a positive/negative growth shock implies a strengthening/weakening of the structural fiscal position.

REFERENCES

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

Cayen, Jean-Philippe, and van Norden, Simon, 2005, The Reliability of Canadian Output-Gap Estimates," *The North American Journal of Economics and Finance,* Elsevier, Vol. 16(3), pp. 373–93, December.

Girouard, N., and C. André, 2005, "Measuring Cyclically-Adjusted Budget Balances for the OECD Countries," OECD Working Paper No. 434 (Paris: OECD Publishing)..

Hughes Hallett, A., R. Kattai, and J. Lewis, 2012, "How Reliable Are Cyclically Adjusted Budget Balances In Real Time?" *Contemporary Economic Policy, Western Economic Association International,* Vol. 30(1), pp. 75–92.

Jonung, Lars, and Martin Larch, 2006, "Improving Fiscal Policy in the EU: The Case for Independent Forecasts," *Economic Policy*, CEPR & CES & MSH, Vol. 21(47), pp. 491–534, 07.

Koske, I., and N. Pain, 2008, "The Usefulness of Output Gaps for Policy Analysis," OECD Economics Department Working Papers No. 621 (Paris: OECD Publishing)..

Ley, Eduardo, and F. Misch, 2013, "Real-time Macro Monitoring and Fiscal Policy," *World Bank Policy Research Working Paper No. 6303.*

Massimiliano, M., and A. Musso, 2011, "The Reliability of Real-time Estimates of the Euro Area Output Gap," *Economic Modelling* Vol. 28, pp. 1842–56.

Orphanides, Athanasios, and van Norden, Simon, 2002, "The Unreliability of Inflation of Output Gap Estimates in Real Time," *The Review of Economics and Statistics,* Vol. 84, pp. 569–83.

———, 2005, "The Reliability of Inflation Forecasts Based on Output Gap Estimates in Real Time," CEPR Discussion Papers 4830.

APPENDIX I. METHODOLOGY FOR DECOMPOSING THE REVISION TO THE CAPB

We define the CAPB as a share of GDP of year *t* as of vintage *v*, as follows:

$$capb_{v,t} = \frac{R(Y_{v,t})}{Y_{v,t}} * \left(\frac{Y^*_{v,t}}{Y_{v,t}}\right) * \left(1 + \frac{Y_{v,t} - Y^*_{v,t}}{Y^*_{v,t}}\right)^{-(\epsilon_R - 1)} - \frac{G^N_{v,t} - I^N_{v,t}}{P_{v,t}Y_{v,t}} * \left(\frac{Y^*_{v,t}}{Y_{v,t}}\right) * \left(1 + \frac{Y_{v,t} - Y^*_{v,t}}{Y^*_{v,t}}\right)^{-(\epsilon_G - 1)}$$

$$= capb_{v,t}(R_{v,t}, G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v,t}, Y^*_{v,t}, \epsilon_R, \epsilon_G)$$

where $R_{v,t}$ represents real revenue and is a function of output, $G^N_{v,t}$ represents the nominal government spending, $I^N_{v,t}$ is the nominal interest spending, and $P_{v,t}$ is the GDP deflator, $Y_{v,t}$ and $Y^*_{v,t}$ are, respectively, the levels of real output and potential output, and ϵ_R and ϵ_G represent the revenue and expenditure elasticities with respect to the output gap. We use the elasticity values for revenue and expenditures used by the EC based on the methodology in Girouard and Andre (2005).

We then decompose the *cross-vintage change* in the CAPB ratio to GDP into the contributions of individual sub-components$(G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v,t}, Y^*_{v,t})$ and a residual cross-term, by allowing one of them to change and keeping all the others constant. For example, in the case of real potential GDP, $Y^*_{v,t}$, we assess its relative contribution to revisions of the CAPB keeping $(G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v,t})$ constant, as follows:

$$\frac{\Delta(\boldsymbol{capb_{v,t}})}{\Delta Y^*_{v,t}} = \boldsymbol{capb_{v+1,t}(R_{v,t}, G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v,t}, Y^*_{v+1,t}, \epsilon_R, \epsilon_G) - capb_{v,t}(R_{v,t}, G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v,t}, Y^*_{v,t}, \epsilon_R, \epsilon_G)} + \text{RESIDUAL}$$

In a similar way, we also calculate the contribution of other subcomponents and a residual term. The only difference is with respect to the contribution of output revisions since the impact of output on the CAPBto GDP ratio is through two channels: the denominator effect as well as indirectly though the cyclical effect on revenues. Therefore, to calculate the contribution of output, we take into account the endogenous impact on revenues. As such, the contribution of output is calculated as:

$$\frac{\Delta(\boldsymbol{capb_{v,t}})}{\Delta Y_{v,t}} = \boldsymbol{capb_{v+1,t}(R_{v+1,t}, G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v+1,t}, Y^*_{v,t}, \epsilon_R, \epsilon_G) - capb_{v,t}(R_{v,t}, G^N_{v,t}, I^N_{v,t}, P_{v,t}, Y_{v,t}, Y^*_{v,t}, \epsilon_R, \epsilon_G) + RESIDUAL}$$

APPENDIX II. THE SIZE AND DISTRIBUTION OF REVISIONS TO OUTPUT AND POTENTIAL OUTPUT OVER ALTERNATIVE HORIZONS

Alternative Revision Horizons

Revisions to output gap, actual and potential output can be computed over a number of horizons. We define these horizons relative to the chosen "final" estimate for these variables, produced in the spring of the 3rd year following the budget year *t*. Therefore the length of the revision horizons range from ½–4 ½ years, with the longest ones including both forecast errors as well as pure backward looking revisions while the short ones, computed starting from vintages *after* year t, include only backward looking revisions. For example, a revision to output gap at the -4.5 y horizon is computed as the "final (F)" Ygap(t+3,spring) minus the "initial (I)" Ygap(t-2,fall).

Fig A2.1 Time horizons for calculation of forecast errors referring to year *t*

Horizon \ Vintage	(t-2,f)	(t-1,s)	(t-1,f)	(t,s)	(t,f)	(t+1,s)	(t+1,f)	(t+2,s)	(t+2,f)	(t+3,s)
-4.5y	I------	------	------	------	------	------	------	------	------	------F
-4.0y		I------	------	------	------	------	------	------	------	------F
-3.5y			I ------	------	------	------	------	------	------	------F
-3.0y				I ------	------	------	------	------	------	------F
-2.5y					I ------	------	------	------	------	------F
-2.0y						I ------	------	------	------	------F
-1.5y							I ------	------	------	------F
-1.0y								I ------	------	------F
-0.5y									I------	------F

Revisions in Output Gap, Output and Potential Output

Output gap estimates are revised continuously. These magnitude of these revisions is nearly 2 ½ –3 ½ percent of potential GDP over 3.5 year horizon—a typical *long-term revision* horizon when macroeconomic assumptions are made for the budget and final output data is available.

Table A2.1 Magnitude and sources of revision in output gap estimates, EU -27, 2004-2012 1/

	-4.5y (t+3,s) - (t-2,f)	-4y (t+3,s) - (t-1,s)	-3.5y (t+3,s) - (t-1,f)	-3y (t+3,s) - (t,s)	-2.5y (t+3,s) - (t,f)	-2y (t+3,s) - (t+1,s)	-1.5y (t+3,s) - (t+1,f)	-1y (t+3,s) - (t+2,s)	-0.5y (t+3,s) - (t+2,f)
Root mean squared error of differences in output gap	4.4	4.2	3.5	3.0	2.5	2.1	1.6	1.2	0.7
Mean of the absolute error of differences in output gap	3.4	3.3	2.6	2.2	1.8	1.5	1.1	0.9	0.5
Root mean squared error of percent change in output level	9.2	8.3	7.1	4.9	3.7	2.8	1.8	1.5	0.4
Mean of the absolute error of percent change in output level	6.9	6.1	5.2	3.3	2.5	1.8	1.1	0.8	0.2
Root mean squared error of percent change in potential output level	5.7	5.4	5.1	3.9	3.5	3.0	2.3	1.8	0.8
Mean of the absolute error of percent change in potential output level	3.7	3.7	3.6	2.7	2.5	2.1	1.6	1.2	0.5

Source: European Commission; and IMF staff calculations.

1/ for 2011 and 2012 the "final (F)" observations are (t+2, s) and, respectively, (t+1,s)

Output and potential output growth revisions follow a similar pattern; at the 3.5 year horizon, the magnitude of revisions ranges between 1–3 ½ percent. Over horizons that compare only ex-post data, the size of revisions is similar between growth and potential growth.

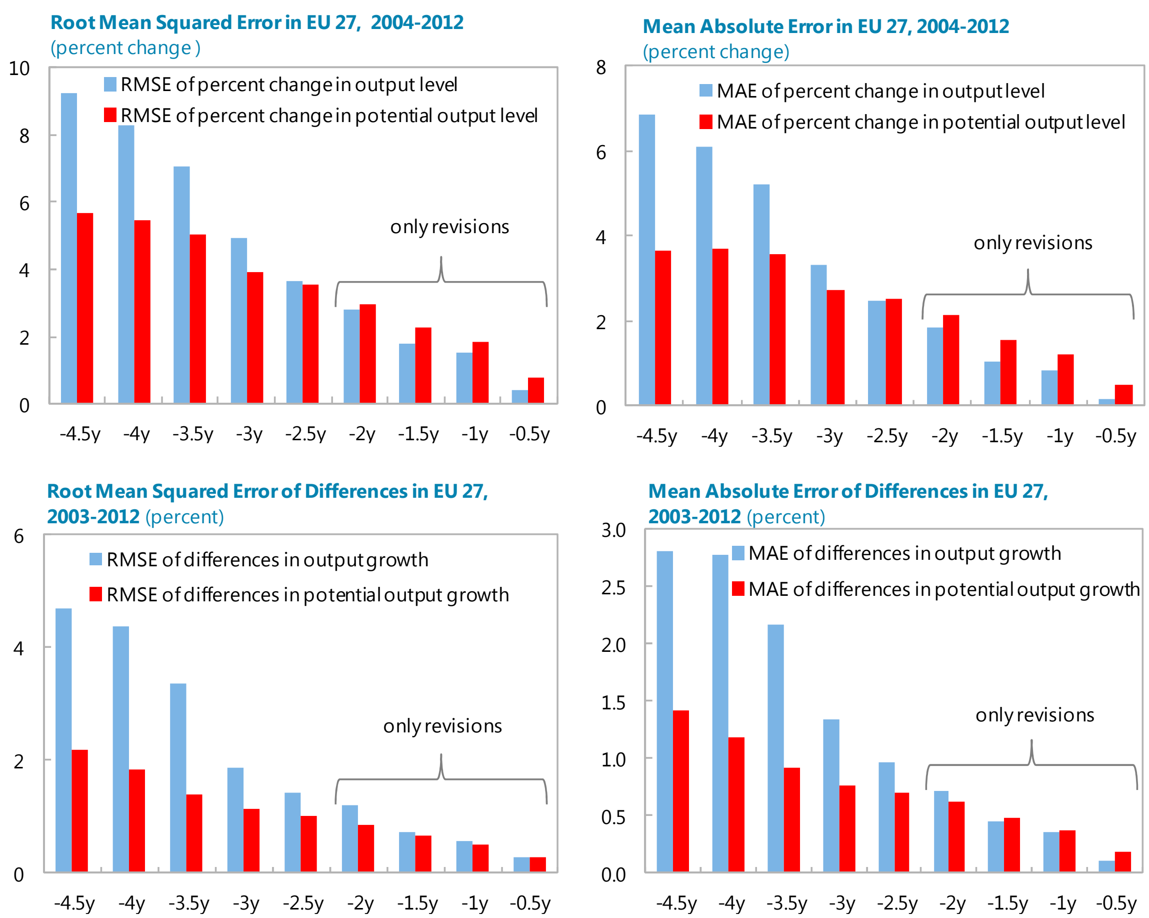
Table A2.2 Magnitude revisions in estimates of growth and potential growth, EU 27, 2003-2012 1/

	-4.5y (t+3,s) - (t-2,f)	-4y (t+3,s) - (t-1,s)	-3.5y (t+3,s) - (t-1,f)	-3y (t+3,s) - (t,s)	-2.5y (t+3,s) - (t,f)	-2y (t+3,s) - (t+1,s)	-1.5y (t+3,s) - (t+1,f)	-1y (t+3,s) - (t+2,s)	-0.5y (t+3,s) - (t+2,f)
Root mean squared error of differences in output growth	4.7	4.4	3.4	1.9	1.4	1.2	0.7	0.6	0.3
Mean of the absolute error of differences in output growth	2.8	2.8	2.2	1.3	1.0	0.7	0.4	0.4	0.1
Root mean squared error of differences in potential output growth	2.2	1.8	1.4	1.2	1.0	0.9	0.7	0.5	0.3
Mean of the absolute error of differences in potential output growth	1.4	1.2	0.9	0.8	0.7	0.6	0.5	0.4	0.2

Source: European Commission; and IMF staff calculations.
1/ for 2011 and 2012 the "final (F)" observations are (t+2, s) and, respectively, (t+1,s)

Revisions in output gap reflected both GDP forecast error as well as revision to estimate of potential GDP. Ex-ante, the revision in GDP is larger whereas ex-post, the revision in potential GDP is larger. This likely reflects, at least in part, the ex-post data revision of GDP levels. The revision in potential would also reflect information on GDP forecasts going forward.

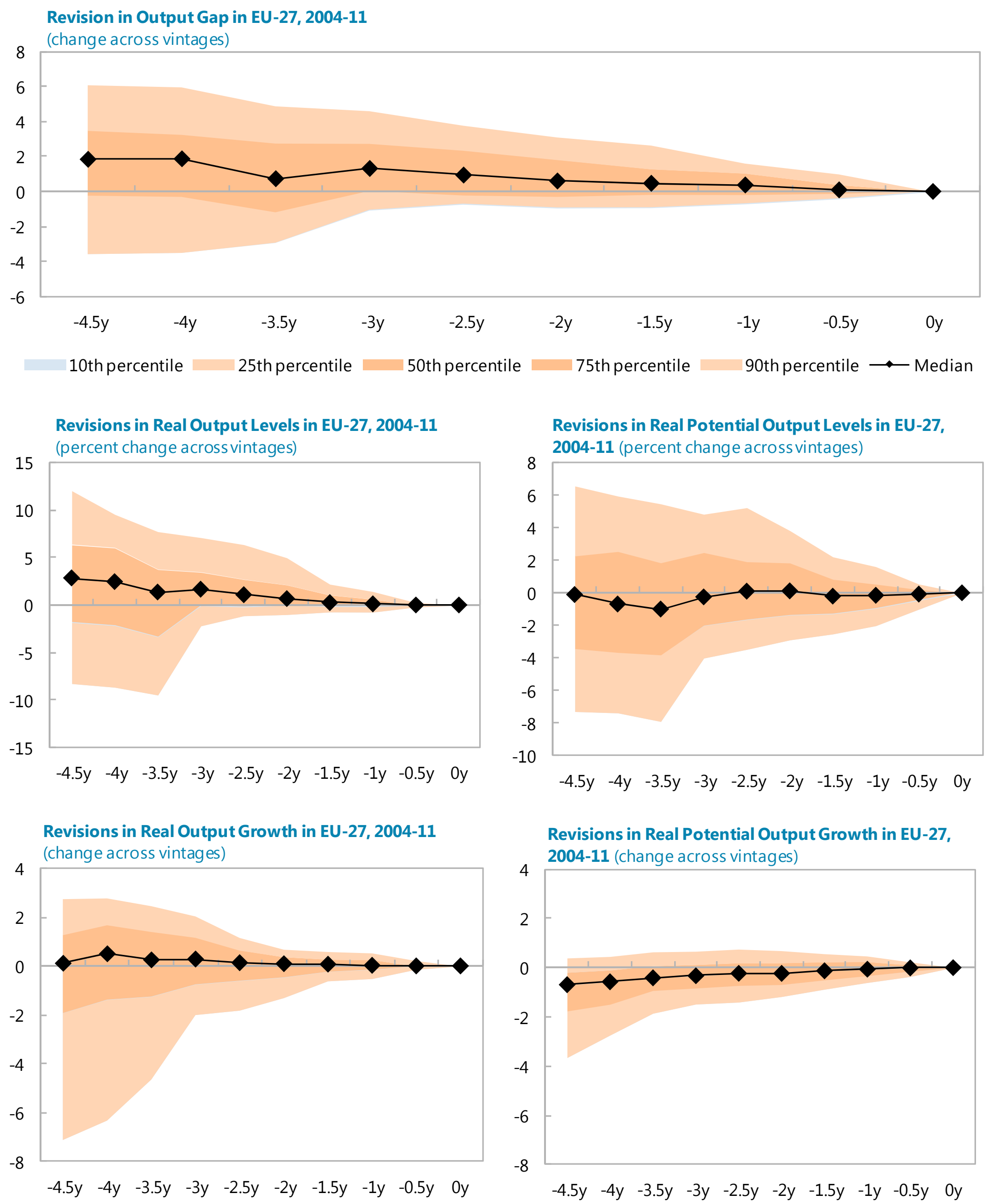
Figure A2.2. The size of revisions diminishes over horizons, with revisions in GDP dominating ex-ante and revisions in potential becoming more important ex-post 1/



Sources: European Commission; and IMF staff calculations.
1/ Calculations contain both ex-post revisions and ex-ante forest errors, except when marked otherwise on the figures.

The distribution of these errors across countries shows that output gap revisions were generally skewed to the positive side (i.e. the output gap was larger following the revision), with a sizable dispersion around the median. The median revision for output was positive (i.e. output was revised upwards) while that for potential output was negative. This suggests that initial estimates of potential were over optimistic for many countries and were subsequently revised down. The large revisions in output reflected forecast errors, and to a lesser extent, ex-post data revisions. Indeed, ex-post revisions were more sizable for potential GDP than output estimates. This ex-post downward revision of potential output likely reflected the uncertainties in long-term output forecasts (which affect the calculation of potential through end-point bias) rather than ex-post output data revisions.

Figure A 2.3. The larger uncertainty in the ex-post revisions of potential compared to actual output appears to be explained mainly by forecast uncertainty



Sources: European Commission; and IMF staff calculations.

APPENDIX III. ESTIMATING THE CROSS-VINTAGE CORRELATION OF OUTPUT AND POTENTIAL OUTPUT GROWTH

In this section, an empirical correlation between short term growth surprises and long term potential growth revisions across a specific horizon is estimated. Given that we are interested in the implications of output gap revisions, arguably, revisions in output and potential output *levels* would be of more direct relevance. However, any correlation analysis would be biased by data revisions. To minimize the impact of such data revisions, growth rates are used rather than levels.[1] Furthermore, instead of using the budget horizon, the following analysis is based on a different horizon, as defined below, in order to better capture other relevant reference points for implementing fiscal policy.

Time horizons for defining revisions

We define a short and long-term horizon for defining growth revisions:

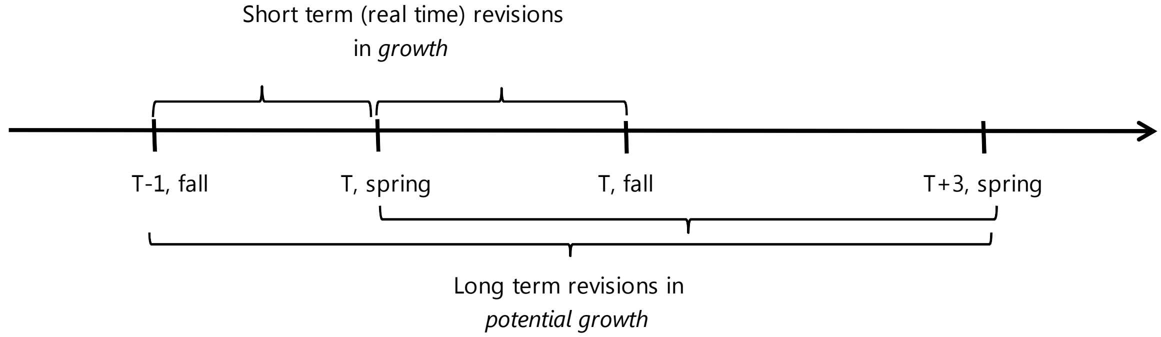
A short-term revision in growth for a given year t is the difference between a new real-time estimate of growth for that year and the estimate produced *one semester* before. The frequency of EC forecasts permits the computation of two real time *revisions* for the current year growth: one between the spring estimate (t,s) and the forecast produced in the fall of the preceding year (t-1,f)—typically the budget preparation period) and another update between the fall (t,f) and spring (t,s) estimates produced during the current year. This revision reflects the growth data that is observed as the budget is being executed.

A long-term revision in growth for a given year is the difference between the *final* estimate of growth for the given year and the real-time estimate produced during the spring of the given year (t,s) (or alternatively, the estimate produced in the fall of the preceding year, the budget preparation period (t-1,f). Our proxy for the *final revision* of the potential growth is represented by the estimate produced by the EC in the spring of the third year following the given year (t+3,s).[2] We chose this horizon as output data is finalized within this three year period.

[1] For completeness, estimations were also done using levels data (for an example, see Table A3.9).

[2] By extending the horizon to three years, we lose the observations in the last two years. Thus, all estimations are based on data from 2003-2010.

Fig A3.1 Short and long term revisions in growth and potential growth (T= budget year)



We illustrate these definitions for budget year 2006 for Spain in the table below.

Table A3.1. Short and long term revisions to growth and potential growth Spain, budget year T = 2006

	2005, fall	2006, spring	2006, fall	...	2009, spring
Growth estimates	3.2	3.1	3.8		3.9
Fall real time revision (2006,spring - 2005,fall)		*-0.1*			
Spring real time revision (2006,fall - 2006,spring)			*0.7*		
Potential growth estimates	3.4	3.6			3.0
Fall long term revision (2009,spring - 2005,fall)		*-0.4*			
Spring long term revision (2009,spring - 2006,spring)			*-0.6*		

Source: European Commision; and IMF staff calculations

Estimation of the empirical correlation

These definitions can be used to assess how much of the short-term growth revision reflects changes in potential growth. To assess how much of the short-term, real-time growth revision could be interpreted as having a permanent component, the long-term revisions in potential growth (the measure for the final, "true," potential growth revision) are regressed on a constant and the short-term, real-time growth revisions, using various country and time effects for both spring and fall growth revisions. In this simple specification, if real-time growth revisions were all permanent, the coefficient for the short-term growth revision would be 1. If the opposite would hold and short-term growth revisions only captured cyclical developments, the coefficient would be 0 instead.

Estimation results

A 1 percentage point short-term output growth revision results on average in some 0.2–0.3 percentage point revision in long run potential growth in the same direction (Table A3.2). This correlation is fairly robust across estimation methods and whether we use data from spring or fall revisions.

Table A3.2. Correlation of long run revisions in potential growth and real time revisions in actual growth, for a given year t, 2003-10

ΔgY*(c,t)= b0 + b1 x ΔgY(c,t) + eps(c,t) 1/ 2/

	Pooled OLS	Country effects	Time effects	Country and time effects
Baseline spring revision 2/				
ΔgY = vintage(t+3, s) - vintage(t,s)*				
ΔgY = vintage(t,f) - vintage(t,s)	0.24***	0.2**	0.3***	0.26***
Nobs	188	188	188	188
Adjusted R2	0.09	0.22	0.37	0.52
Baseline fall revision 3/				
ΔgY = vintage(t+3, s) - vintage(t-1,f)*				
ΔgY = vintage(t,s) - vintage(t-1,f)	0.32***	0.28***	0.54***	0.39***
Nobs	186	186	186	186
Adjusted R2	0.31	0.42	0.52	0.61

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).

2/ For simplicity and ease of comparison with the baseline estimation, tables A3.2 to A3.8 list the sample of budget years as 2003-2010, given that the latest forecast vintage in our dataset is 2013,s (hence the latest budget year for which a t+3,s observation can be computed is 2010). However, for the shorter horizons in some of the robustness checks (e.g. semi-annual rolling, T A3.7) more budget years can be included (e.g. through 2012).

3/ Refers to a real time revision of actual growth between the fall and spring of the budget year.

4/ Real time revision of actual growth between the spring vintage of the budget year and the fall vintage of the budget preparation year.

As a robustness check, we test whether the results differ between EU New Member States (Table A3.3) and Old Member States (Table A3.4). This should help as estimation of potential growth is considered especially problematic in rapidly converging economies. Results suggest that this does not seem to be the case as our baseline estimates are aligned between the two subsamples.

Table A3.3 Correlation of long run revisions in potential growth and real time revisions in actual growth, for a given year t, t=2003:2010, EU New Member

$\Delta gY^*(c,t) = b0 + b1 \times \Delta gY(c,t) + eps(c,t)$ 1/

	Pooled OLS	Country effects	Time effects	Country and time effects
Baseline spring revision 2/				
ΔgY = vintage(t+3, s) - vintage(t,s)*				
ΔgY = vintage(t,f) - vintage(t,s)	0.23**	0.19*	0.34***	0.3***
Nobs	68	68	68	68
Adjusted R2	0.07	0.18	0.42	0.58
Baseline fall revision 3/				
ΔgY = vintage(t+3, s) - vintage(t-1,f)*				
ΔgY = vintage(t,s) - vintage(t-1,f)	0.34***	0.3***	0.54***	0.35***
Nobs	66	66	66	66
Adjusted R2	0.32	0.40	0.53	0.63

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).
2/ Refers to a real time revision of actual growth between the fall and spring of the budget year.
3/ Refers to a real time revision of actual growth between the spring of the budget year and the fall of the budget preparation year.

Table A3.4. Correlation of long run revisions in potential growth and real time revisions in actual growth, for a given year t, t=2003:2010, EU Old Member States

$\Delta gY^*(c,t) = b0 + b1 \times \Delta gY(c,t) + eps(c,t)$ 1/

	Pooled OLS	Country effects	Time effects	Country and time effects
Baseline spring revision 2/				
ΔgY = vintage(t+3, s) - vintage(t,s)*				
ΔgY = vintage(t,f) - vintage(t,s)	0.24**	0.23*	0.36***	0.36***
Nobs	120	120	120	120
Adjusted R2	0.10	0.14	0.41	0.49
Baseline fall revision 3/				
ΔgY = vintage(t+3, s) - vintage(t-1,f)*				
ΔgY = vintage(t,s) - vintage(t-1,f)	0.25***	0.23***	0.5***	0.42***
Nobs	120	120	120	120
Adjusted R2	0.26	0.30	0.48	0.53

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).
2/ Refers to a real time revision of actual growth between the fall and spring of the budget year.
3/ Refers to a real time revision of actual growth between the spring of the budget year and the fall of the budget preparation year.

A number of alternative sample sizes and revision horizons were also tested to verify the robustness of the results One potential objection to the baseline specification is that the choice of the (t+3, spring) vintage as the final estimate for potential GDP growth may in fact not represent the true outcome as even this estimate could be subject to subsequent revisions. An alternative is to compute potential growth revisions relative to the *latest available* estimate (the 2013, spring forecast vintage, in our sample) of potential growth for budget year *t* (Table A3.5). Symmetrical as well as shorter horizons for both growth and potential growth revisions were also used for alternative specifications (Tables A3.6 and A3.7). To try to isolate the effect of the 2008-09 crisis, the data subsample was changed to include only non-crisis years (Table A3.8). Results remained broadly stable across this array of specifications.

Table A3.5. Correlation of long run revisions in potential growth and real time revisions in actual growth, for a given year t, 2003-10, latest available observation

ΔgY*(c,t)= b0 + b1 x ΔgY(c,t) + eps(c,t) 1/

	Pooled OLS	Country effects	Time effects	Country and time effects
Alternative spring revision 2/				
ΔgY = vintage(2013, s) - vintage(t,s)*				
ΔgY = vintage(t,f) - vintage(t,s)	0.26***	0.25***	0.3***	0.29***
Nobs	242	242	242	242
Adjusted R2	0.13	0.23	0.31	0.43
Alternative fall revision 3/				
ΔgY = vintage(2013, s) - vintage(t-1,f)*				
ΔgY = vintage(t,s) - vintage(t-1,f)	0.35***	0.32***	0.52***	0.44***
Nobs	240	240	240	240
Adjusted R2	0.35	0.41	0.49	0.54

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0). The final estimate of potential growth for a given year t is taken to be the latest currently available estimate, namely the 2013, spring vintage.

2/ Refers to a real time revision of actual growth between the fall and spring of the budget year.

3/ Real time revision of actual growth between the spring vintage of the budget year and the fall vintage of the budget preparation year.

While the main motivation for using real time growth revisions was to inform a rule of thumb that could be useful for real time evaluations of current fiscal adjustment, the empirical correlation between growth and potential growth can also be estimated across alternative horizons reflecting different decision points for assesing the implementation of fiscal policy. In particular, a *budget preparation horizon* can be defined as the revision from the time of buget prepartion (t-1,f) to the "final" estimate of both growth and potential growth 3 years later (t+3,s). Alternatively, the correlation can be estimated over the *budget evaluation horizon* as defined in section II. Finally, one can also compare the growth and potential growth figures between the "final" estimate at (t+3,s) and (t,s), designated as the *budget implementation horizon* (Table A3.6).The results from estimating correlations also over shorter term (semi-annual) horizons suggest, as expected, that the degree to which growth revisions have permanent effects diminishes over time (Table A3.7).

Table A3.6. Correlation of revisions in potential growth and actual growth, for a given year t, 2003-10, alternative horizons

ΔgY*(c,t)= b0 + b1 x ΔgY(c,t) + eps(c,t) 1/

	Pooled OLS	Country effects	Time effects	Country and time effects
Budget preparation horizon				
ΔgY, ΔgY = vintage(t+3,s) - vintage(t-1,f)*	0.26***	0.23***	0.34***	0.29***
Nobs	186	186	186	186
Adjusted R2	0.49	0.61	0.65	0.75
Budget evaluation horizon				
ΔgY, ΔgY = vintage(t+1,s) - vintage(t-1,f)*	0.31***	0.3***	0.35***	0.35***
Nobs	240	240	240	240
Adjusted R2	0.73	0.73	0.81	0.82
Budget implementation horizon				
ΔgY, ΔgY = vintage(t+3,s) - vintage(t,s)*	0.28***	0.24***	0.33***	0.29***
Nobs	188	188	188	188
Adjusted R2	0.22	0.34	0.46	0.59

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).

Table A3.7. Correlation of revisions in potential growth and actual growth, for a given year t, t=2003:2010, rolling semi-annual horizons

ΔgY*(c,t)= b0 + b1 x ΔgY(c,t) + eps(c,t) 1/

	Pooled OLS	Country effects	Time effects	Country and time
vintage(t,s) - vintage(t-1,f)	0.29***	0.29***	0.29***	0.3***
Nobs	240	240	240	240
Adjusted R2	0.79	0.79	0.79	0.79
vintage(t,f) - vintage(t,s)	0.29***	0.3***	0.31***	0.32***
Nobs	242	242	242	242
Adjusted R2	0.54	0.54	0.60	0.60
vintage(t+1,s) - vintage(t,f)	0.39***	0.43***	0.26***	0.3***
Nobs	252	252	252	252
Adjusted R2	0.36	0.40	0.47	0.49
vintage(t+1,f) - vintage(t+1,s)	0.27***	0.29***	0.24***	0.27***
Nobs	227	227	227	227
Adjusted R2	0.05	-0.02	0.20	0.14
vintage(t+2,s) - vintage(t+1,f)	0.11	0.21	0.09	0.18**
Nobs	237	237	237	237
Adjusted R2	0.00	-0.01	0.38	0.41
vintage(t+2,f) - vintage(t+2,s)	0.18**	0.19**	0.19**	0.2**
Nobs	212	212	212	212
Adjusted R2	0.04	-0.01	0.20	0.16
vintage(t+3,s) - vintage(t+2,f)	0.11**	0.08	0.07**	0.03
Nobs	212	212	212	212
Adjusted R2	0.01	-0.07	0.45	0.43

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).

Our full data sample may suffer from structural breaks due to the inclusion of crisis years that saw significant revisions to potential growth compared to original estimates. One possible robustness check is to estimate the empirical correlation over a subsample of observations excluding crisis years 2008-09 and at short term (semi annual) horizons (Table A 3.8). These estimates support our baseline results, although they are not exactly comparable as our baseline estimation refers to the correlation of short term growth surprises with permanent (long term) changes in potential growth.

Table A3.8. Correlation of revisions in potential growth and actual growth, for a given year t, t=2003-07,2010, rolling semi-annual horizons (non-crisis)

ΔgY*(c,t)= b0 + b1 x ΔgY(c,t) + eps(c,t) 1/

	Pooled OLS	Country effects	Time effects	Country and time effects
vintage(t,s) - vintage(t-1,f)	0.3***	0.27***	0.3***	0.27***
Nobs	186	186	186	186
Adjusted R2	0.42	0.43	0.41	0.42
vintage(t,f) - vintage(t,s)	0.27***	0.28***	0.3***	0.34***
Nobs	188	188	188	188
Adjusted R2	0.39	0.34	0.45	0.42
vintage(t+1,s) - vintage(t,f)	0.27***	0.31***	0.24***	0.26***
Nobs	198	198	198	198
Adjusted R2	0.17	0.14	0.20	0.17
vintage(t+1,f) - vintage(t+1,s)	0.29***	0.3**	0.26***	0.27***
Nobs	173	173	173	173
Adjusted R2	0.07	-0.04	0.24	0.15
vintage(t+2,s) - vintage(t+1,f)	0.15	0.22	0.13	0.19*
Nobs	183	183	183	183
Adjusted R2	0.01	-0.04	0.41	0.43
vintage(t+2,f) - vintage(t+2,s)	0.15	0.15	0.13	0.14
Nobs	158	158	158	158
Adjusted R2	0.02	-0.03	0.17	0.15
vintage(t+3,s) - vintage(t+2,f)	0.13	0.16	0.04	0.05
Nobs	158	158	158	158
Adjusted R2	0.00	-0.12	0.47	0.44

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).

Finally, to address the possibility of omitted variable bias in the case where current year growth surprises affect the projections for subsequent years, we have also re-estimated the baseline model by including real time growth revisions for subsequent years in addition to year *t*. Our main results remain robust in this additional specification (Table 3.9).

Table A3.9 Correlation of long run revisions in potential growth and real time revisions in actual growth, for a given year t , controlling for multi year growth revisions, 2003-10

ΔgY*(c,t)= b0 + b1 x ΔgY(c,t) + b2 x ΔgY(c,t+1) + b3 x ΔgY(c,t+2) + eps(c,t) 1/ 2/

	Pooled OLS	Country effects	Time effects	Country and time effects
Baseline spring revision 3/				
ΔgY(t) = vintage(t+3, s) - vintage(t,s)*				
ΔgY(t) = vintage(t,f) - vintage(t,s)	0.28***	0.29***	0.32***	0.32***
ΔgY(t+1) = vintage(t,f) - vintage(t,s)	0.41***	0.41***	0.29***	0.3***
ΔgY(t+2) = vintage(t,f) - vintage(t,s)	0.25***	0.26***	0.23***	0.24***
Nobs	188	188	188	188
Adjusted R2	0.49	0.62	0.57	0.72
Baseline fall revision 4/				
ΔgY = vintage(t+3, s) - vintage(t-1,f)*				
ΔgY(t) = vintage(t,s) - vintage(t-1,f)	0.4***	0.36***	0.53***	0.49***
ΔgY(t+1) = vintage(t,s) - vintage(t-1,f)	0.3***	0.27***	0.49***	0.45***
ΔgY(t+2) = vintage(t,s) - vintage(t-1,f)	0.23***	0.2***	0.28***	0.23***
Nobs	186	186	186	186
Adjusted R2	0.68	0.72	0.76	0.78

1/ ΔgY* = gY*(c,t,vintage 1) - gY*(c,t,vintage 0); ΔgY = gY(c,t,vintage 1) - gY(c,t,vintage 0).

2/ For simplicity and ease of comparison with the baseline estimation, tables A3.2 to A3.8 list the sample of budget years as 2003-2010, given that the latest forecast vintage in our dataset is 2013,s (hence the latest budget year for which a t+3,s observation can be computed is 2010). However, for the shorter horizons in some of the robustness checks (e.g. semi-annual rolling, T A3.7) more budget years can be included (e.g. through 2012).

3/ Refers to a real time revision of actual growth between the fall and spring of the budget year.

4/ Real time revision of actual growth between the spring vintage of the budget year and the fall vintage of the budget preparation year.

Alternative Specification for Estimating the Cross-Vintage Correlation of Output and Potential Output Levels

The following result replicates the regressions in table A3.2 results using log differences between levels of output and, respectively, potential output across vintages (Table A3.10).

Table A3.10. Correlation of long run revisions in potential output and real time revisions in actual output, for a given year t, 2004-09

ΔY*(c,t)= b0 + b1 x ΔY(c,t) + eps(c,t) 1/

	Pooled OLS	Country effects	Time effects	Country and time
Baseline spring revision 2/				
ΔY = vintage(t+3, s) - vintage(t,s)*				
ΔY = vintage(t,f) - vintage(t,s)	0.77***	0.66***	0.74***	0.63***
Nobs	146	146	146	146
Adjusted R2	0.16	0.19	0.52	0.63
Baseline fall revision 3/				
ΔY = vintage(t+3, s) - vintage(t-1,f)*				
ΔY = vintage(t,s) - vintage(t-1,f)	0.74***	0.63***	0.95***	0.67***
Nobs	144	144	144	144
Adjusted R2	0.24	0.27	0.56	0.63

1/ ΔY* = ln(Y*(c,t,vintage 1)) - ln(Y*(c,t,vintage 0)); ΔY = ln(Y(c,t,vintage 1)) - ln(gY(c,t,vintage 0))

2/ Refers to a real time revision of actual growth between the fall and spring of the budget year.

3/ Real time revision of actual growth between the spring vintage of the budget year and the fall vintage of the budget preparation year.