

THE DESIGN AND IMPLEMENTATION OF LARGE FISCAL ADJUSTMENT PLANS IN THE EUROPEAN UNION¹

S. Ali Abbas, Fuad Hasanov, Paolo Mauro and Junhyung Park

Fiscal Affairs Department, International Monetary Fund, Washington, DC

ABSTRACT

We study the design and implementation of large fiscal adjustment plans using a novel approach. Previous studies have focused on successful fiscal adjustments identified on the basis of outcomes. Instead, we analyze large fiscal adjustment plans and compares them with outcomes, thereby capturing lessons not only from successes but also failures, and avoiding sample selection/survivorship bias. In this paper, we analyze all large fiscal adjustment plans put forward by European Union countries in 1991–2008. On *design*, we find that large planned adjustments envisaged a greater role for expenditure cuts. Only a third of the plans provided for revenue increases, but most of these were not anchored in tangible tax policy measures. Plans were generally based on assumptions of improving macroeconomic conditions, but growth projections were not systematically optimistic relative to independent forecasts available at the time. On *implementation*, we find that three-fourths of the planned adjustment was realized on average, with sizable expenditure over-runs partly offset by revenue over-performance. Economic growth surprises are found to be key: one percentage point higher-than-planned growth improves the overall balance by ½ percentage point of GDP. There is some evidence that policymakers use growth and other surprises (such as revisions to initial-year deficits) opportunistically: partially “spending” favorable surprises, but permitting unfavorable ones to result in deviations from targets.

¹ Abbas, Hasanov, and Mauro are in the International Monetary Fund’s Fiscal Affairs Department. Park is now a Ph.D. student at UCLA. We are grateful to Roel Beetsma for valuable comments and suggestions, to Katia Chen, Alica Dzelilovic, and Patricia Quiros for collecting the plans from various archival sources. The opinions expressed are those of the authors and should not be attributed to the International Monetary Fund. This paper is part of a broader effort on [Chipping Away at Public Debt—Sources of Failure and Keys to Success in Fiscal Adjustment](#) (Wiley, 2011), also summarized in the publicly available [video](#) by the same title.

I. INTRODUCTION

The global financial and economic crisis that began in 2008 has resulted in the largest increase in government debt across advanced economies since the Second World War. As many governments prepare medium-term adjustment plans, a key policy question—currently generating substantial public debate—is how to ensure successful fiscal adjustment in the period ahead. Systematic analysis of past adjustment episodes may inform the current policy debate.

Previous studies have identified successful fiscal adjustment episodes, for the most part, on the basis of the largest observed improvements in the government debt or the overall fiscal balance. In this paper we take an alternative, complementary approach. Starting from *ex-ante* planned fiscal adjustments in Europe over the last two decades, rather than actual *ex-post* fiscal outturns, we identify large fiscal adjustment plans based on intended reductions in headline and cyclically-adjusted budget deficits. The empirical analysis then tracks ex-post outcomes compared with ex-ante plans and attempts to explain the deviations. For instance, we look at the extent to which macroeconomic variables such as economic growth deviated from those projected at the time when the plans were drawn up.

Studies based upon the traditional approach in which large fiscal adjustment are defined on the basis of ex-post outcomes have found interesting and policy relevant results for both advanced economies (e.g.; Alesina and Ardagna, 1998, 2009; von Hagen et al., 2010) and broader samples of countries (Giavazzi et al, 2000; Gupta et al., 2005; Baldacci et al. 2006). The focus has been on whether fiscal adjustments are longer lasting and more successful in eliciting a non-Keynesian growth response when they rely on expenditure cuts rather than on tax hikes. These studies have emphasized the possibility that sustained spending-led reductions (especially cuts in public employment, public wages and transfers) are more successful in terms of their debt and growth aftereffects, including due to favorable supply-side effects.² Factors such as the presence of fiscal stress at the start of the episode, political support, conciliatory trade unions, and multi-year frameworks are also seen to facilitate success.

By focusing on ex-ante plans in the context of a systematic statistical analysis, however, this paper connects also with more recent studies that analyze the credibility and implementation record of announced budgets using “real-time” fiscal data. This literature has also focused on countries in the European Union, because of the availability of relatively standardized convergence or Stability and Growth Programs (SGP) (Strauch et al., 2004; Beetsma and Giuliodori, 2008; Cimadomo, 2008; Beetsma et al., 2009; and von Hagen, 2010). The focus of these studies has been on identifying the determinants of the size and direction of planned fiscal policy changes (such as starting fiscal position, stage of economic and political cycle) and of the factors affecting their implementation across all budgetary plans. These studies have argued that (i) implementation of budgets has generally been weak in the EU; (ii) implementation has been weaker for more ambitious fiscal adjustment plans, which are typically drawn up to signal competence ahead of elections, and/or when the initial deficit is

² Baldacci et al (2006) suggest, for emerging markets, that expenditure reductions are not sufficient, especially when initial revenue ratios are low; and that backloaded adjustments, in which the adjustment is spread over time, work better than frontloaded adjustments.

high;³ and (iii) implementation has been aided by positive growth surprises and strong fiscal institutions.

As our motivation in conducting this empirical analysis is similar to that of the “large fiscal adjustments” literature, but appealing methodological features can be drawn from the “fiscal policy credibility” literature, we bring together helpful elements from both approaches. We focus on episodes of *large* (defined below) planned fiscal adjustments, whereas previous studies in the “fiscal policy credibility” literature considered together not only attempts at large and small fiscal consolidations but also fiscal expansions—episodes that we consider to be fundamentally different in nature. We evaluate large fiscal adjustment plans over a *medium-term* (three-year) horizon, whereas the “fiscal policy credibility” literature worked primarily at the *annual* frequency.⁴ Our focus on large medium-term plans is motivated by the current policy context where many countries are preparing multi-year strategies for large fiscal consolidation to bring debt down from historic post-crisis levels. Moreover, adjustments planned over several years may better capture medium-term fiscal policy goals than a one-year plan whose construction and outcomes are likely contaminated by various short-term factors. Finally, we work with a longer sample period—beginning in 1991—than previous studies did.

The main findings of our empirical analysis on the design and implementation of plans are as follows:

- On *design*, we find that large planned adjustments typically envisaged a greater role for expenditure cuts than tax increases (indeed in a majority of cases, tax *reductions* were planned). Of the one-third of plans that envisaged revenue increases, less than half were anchored in tangible tax policy measures, but in most of these cases, the measures were implemented and resulted in significant revenue increases that were generally sustained beyond the plan horizon. Most plans were based on assumptions of improving macroeconomic conditions, although growth projections were not found to be optimistic relative to IMF staff forecasts published in the *World Economic Outlook* at the time. Plans were noticeably more ambitious when the initial deficit was above the 3 percent of GDP Maastricht criteria, and in the run up to European monetary integration (EMU).
- On *implementation*, we find that, on average, three fourths of the planned adjustment were realized, but that the composition of adjustment often differed from plans, with revenue overperformance offsetting expenditure overruns. Implementation was not adversely affected by plan ambition. There is only partial evidence that carrots such

³ The level of initial debt and “structural” balance do not affect ambition in the overall fiscal balance. In other words, the *corrective* arm of the SGP (excessive deficit procedure) has been more successful in inducing fiscal corrections than the *preventive* (neutral fiscal stance over the cycle) and *dissuasive* arms (debt/GDP should not exceed 60 percent).

⁴ Previous studies have compared the planned adjustment in $t+1$ with actual adjustment in $t+1$, or the planned “year-on-year” adjustment in $t+2$ with the actual “year-on-year” adjustment in $t+2$. We compare 3-year cumulative planned adjustments with three-year cumulative outcomes.

as EMU accession, and sticks such as the European Union’s Excessive Deficit Procedure (EDP) strengthened adherence to plans; the evidence does not survive in regressions where other factors, especially growth surprises, are controlled for. Growth surprises, in fact, turn out to be the most important factor underlying deviations of plans from outcomes: for every one percentage point in higher-than-expected growth, the overall balance improves by 0.5 percent of GDP. Base effects (surprise revisions in the initial fiscal balance) are followed by small changes in end-point fiscal balance targets and thus imply large effects on implementation measured by outcomes versus plans with respect to the improvement in the fiscal balance. Implementation is greater in the presence of stronger national fiscal rules, less fractionalized parliaments, and larger improvements in government stability.

In the remainder of this paper, we discuss data, methodology, and results. Section 2 discusses the data and methodological issues, including the construction of a sample of 66 large adjustment plans in the EU from the universe of convergence programs and SGPs published during 1991–2008. Section 3 reports summary statistics on the design and implementation of the plans in this sample, and provides descriptive statistics and plots on major variables of interest. Section 4 presents our main econometric results on potential drivers of plan implementation such as initial macro-fiscal conditions, the ambition of the planned consolidation, macroeconomic surprises, policy reaction to news, and institutional and political variables. Section 5 concludes with key policy issues.

II. DATA, SAMPLE SELECTION, AND METHODOLOGY

Our starting point is the universe consisting of 229 convergence programs and SGPs prepared by 25 EU countries during 1991–2008: of these, the 175 post-1998 plans were drawn from the European Commission’s website; the remaining 54 pre-1998 plans were sourced from country authorities, libraries and the IMF’s own archives. As the time span of plans published in 2006 and later overlaps significantly with the global financial crisis—whose impact would unduly dominate and skew the results of any systematic statistical analysis of this type—we focus on the 178 published plans before end-2005. For these plans we extracted all available time series information on key macro-fiscal variables—notably the estimates for year t (the year of the plan’s publication), reported outturns for year $t-1$, and projections for years $t+1$, $t+2$, and $t+3$. The fiscal variables include headline balances, revenues, primary expenditures, interest payments and public debt levels. Macroeconomic variables comprise real and nominal GDP growth rates and levels, inflation and interest rates. We also recovered plan information on cyclically adjusted balances for the plans for which it was available (about one-half of the plans), and supplemented that information with our own computations of the cyclically-adjusted balance, which we undertook using carefully constructed real-time output gaps—that is, using only the information that would have been available to contemporary observers (see Appendix 1). We drew ex-post data for the same variables from the European Commission’s AMECO database.

Starting from the 178 pre-2006 plans (corresponding to all shaded cells in Table 1 below), and defining as “large planned fiscal adjustments” those that envisaged a cumulative adjustment in the headline fiscal balance of more than 1 percent of GDP in the following three years, we identified a total of 100 large fiscal adjustment plans (shown in dark gray

shaded cells in Table 1 below). Finally, for each country, plans with more than two overlapping years were dropped, yielding our sample of 66 large planned headline balance improvements.⁵

Table 1. Selection of 66 Large Fiscal Adjustment Plans from the Universe of SCPs

	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
Austria						•				•					•
Belgium		•				•			•						
Cyprus														•	•
Czech Republic															•
Germany			•			•		•			•		•		•
Denmark				•		•		•							
Estonia															
Greece		•		•				•		•		•		•	
Spain	•						•		•					•	
Finland					•		•		•						
France			•			•		•			•		•		•
Hungary															•
Ireland															
Italy	•						•		•		•		•		•
Lithuania														•	
Luxembourg															•
Latvia															
Malta														•	•
The Netherlands	•			•									•		
Poland														•	
Portugal	•		•				•		•		•		•		•
Sweden					•		•						•		
Slovakia															•
Slovenia														•	
United Kingdom			•		•		•						•		•

Universe of plans
 Greater than 1% adjustment
• Large Non-overlapping Plans

As can be gleaned from Table 1, 27 of the 66 plans were published prior to 1998. As the ECOFIN website only posts plans published from 1998 onward, these earlier plans have not been looked at in detail in the fiscal policy credibility literature. Yet, they are of great interest, particularly because they include countries/years characterized by large fiscal adjustments, including in the run-up to EMU. Of these, 17 plans were prepared during 1992–96 and aimed at delivering deficits below 3 percent of GDP by 1997 (the test year for the first round of EMU accession).

Having identified the plans we'll work with, we now discuss two key steps involved in defining fiscal adjustment and tracking implementation:

1. *Defining fiscal adjustment:* We define the ex-ante planned adjustment as the projected percentage point increase in the fiscal balance-to-GDP ratio between years t and

⁵ We also conducted a parallel exercise, where the sorting of plans was done based on planned improvement in the *structural primary balance-to-potential GDP* ratio. For this purpose, we used a threshold improvement of 0.5 percentage points of potential GDP (close to the median improvement in the ratio of the structural primary balance to potential GDP) to define large planned adjustments. This gave us 74 plans, of which 48 were non-overlapping (as defined above). The results for this sample were broadly similar to those described in this paper, and are not reported for the sake of brevity.

$t+3$ as reported in the year t plan. The corresponding ex-post adjustment is defined as the actual percentage point increase in the fiscal balance-to-GDP ratio between years t and $t+3$ as reported in the final data (the 2009 AMECO vintage).⁶ In addition to the overall fiscal balance and its revenue and expenditure components, we also track adjustment in the cyclically-adjusted primary balance, sub-divided into cyclically-adjusted revenues and primary expenditures.

2. *Measuring implementation:* We use two measures of plan implementation: (a) an implementation *ratio* which, for each plan, scales the *actual* observed three-year cumulative improvement in the fiscal balance-to-GDP ratio to the corresponding planned improvement; (b) an implementation *error* which, for each plan, subtracts the planned three-year cumulative improvement in the fiscal balance-to-GDP ratio from the corresponding actual improvement. Perfect implementation would be implied by an implementation ratio of 1 or, equivalently, an implementation error of zero. Of course, over-performance of plans (implementation ratio greater than 1, or implementation errors greater than zero) is also feasible.

An advantage of using the implementation *ratio* is that it appropriately scales actual performance to the plan's fiscal adjustment ambition (planned improvement in the fiscal balance) and is thus comparable across plans in a non-regression framework (Section 3), i.e., where ambition is not otherwise controlled for. An advantage of using the implementation *error* is the ease of interpretation it affords in a regression framework, as its units are percentage points of GDP. We, therefore, use this measure in our formal regressions (Section 4), where ambition is appropriately controlled for.

III. DESCRIPTIVE ANALYSIS

In this section, we present summary statistics to shed light on key design aspects of the plans in our sample, aggregate statistics on the plans' implementation, and some descriptive analysis on potential drivers of implementation.

A. Plan Design

A quick look at the share of the plans in our sample that display certain features reveals a wide mix, making for interesting comparisons, and begins to provide a sense of the results which will be confirmed through more systematic analysis in the next sections.

As noted earlier, two-fifths of the plans are pre-1998. Over half of the plans were prepared in the context of initial deficits or debts above the Maastricht-prescribed ceilings of 3 and 60 percent of GDP, respectively. Moreover, although this would not always have been known at the time the plans were drawn up, a change in government during the three-year

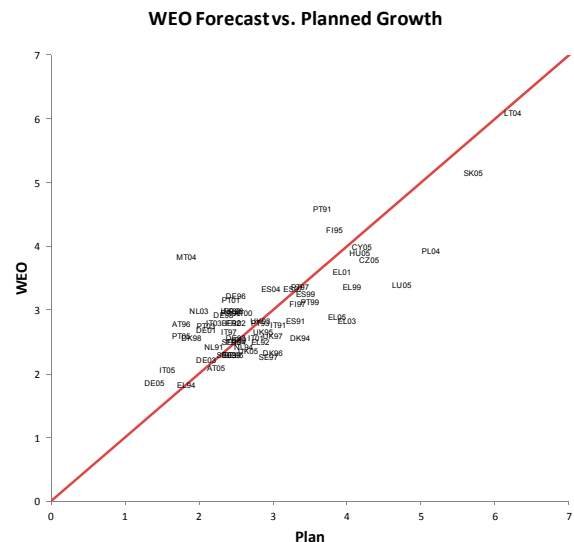
⁶ By extracting all ex-ante information from a single plan document, and all ex-post information from a single AMECO vintage, we limit the impact of base effects (i.e., revisions to initial deficit data) and changes in accounting definitions *within* the ex-ante or the ex-post data (though such differences may emerge *between* ex-ante and ex-post data).

adjustment horizon occurred in about half the cases. Less than a quarter of the plans envisaged frontloaded adjustment (i.e., more than half of the adjustment occurring in the first year).

Various Plan Features	Proportion of 66 Plans (in percent)
Dated pre-1998	41
Where initial deficit >3 percent of GDP	56
Where initial debt >60 percent of GDP	59
Prepared in run up to EMU	23
During which there was a change in government	47
With >50% adjustment in year 1	23
Projecting output gap increase	77
With >50% adjustment from spending side	87
Stipulating revenue effort backed by measures	17

Three-quarters of the plans in our sample projected an improvement in macroeconomic conditions as measured by the output gap (i.e., the percentage deviation of actual real output from potential) over the plan horizon. This is not surprising given that one of policymakers' key objectives is to foster economic growth and that EMU entry or broader European integration bore the promise of economic gains. Figure 1 shows that real GDP growth rate assumptions embedded in the plans were not overly optimistic, on average, and that, for the most part, they were close to IMF staff projections published in the contemporary issues of the *World Economic Outlook* (most observations are close to the 45 degree line).⁷ Thus, ex-post growth surprises relative to plan must be ascribed to true macroeconomic uncertainty rather than any systematic projection bias (see also IMF, 2008).

Figure 1: WEO Growth Forecasts Vs Plan Growth Rates (Percent per annum, annualized from three-year cumulative)



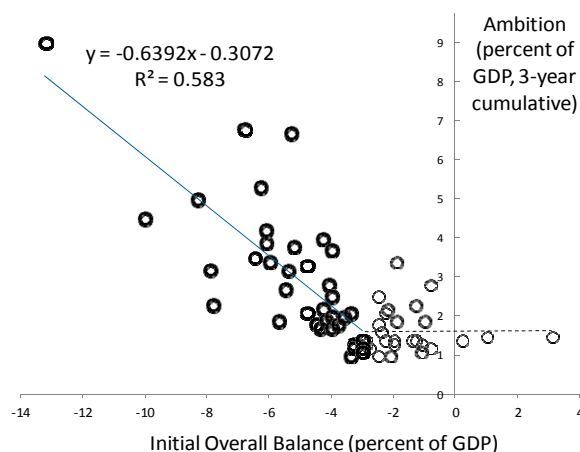
Planned adjustments were more ambitious the higher the initial deficit, although the relationship appears to be more relevant at higher deficit levels. Figure 2 reports this bivariate association separately for plans formulated when the initial deficit was larger than the EMU-prescribed limit of 3 percent of GDP (dark circles and solid regression line) and for plans formulated when the initial deficit was already below the 3 percent limit (fainter circles and dashed regression line). Greater ambition was significantly associated with larger initial deficits (with a slope coefficient above 0.6), but only when these were above 3 percent.

⁷ A *t*-test on the deviations of economic growth rates assumed under the plans from WEO growth rates published around the same time did not reject the null hypothesis that the distribution of these deviations was centered on zero. A test of forecast unbiasedness further confirmed that plan growth forecasts were unbiased relative to “actual” growth rates.

When initial deficits were below 3 percent, the association with plans' ambition is essentially nil. These general patterns in policymakers' planning behavior are somewhat reassuring, because a necessary condition for fiscal sustainability is that fiscal performance be improved in response to a worsening in the initial fiscal position.⁸

The plans' composition in terms of the expenditure-revenue measure mix overwhelmingly favored expenditure cuts over revenue increases. Almost 90 percent of the plans envisaged that more than half of the adjustment would come from spending cuts (as shares of GDP). Moreover, almost two-thirds of the plans envisaged that revenues would be cut as a share of GDP, thus requiring cuts in expenditure ratios in excess of the total adjustment in the fiscal balance.

Figure 2: Plan Ambition vs. Initial Balance



Revenue and Expenditure Composition of 66 Plans (in percent)				
		Expenditure		
		Increase	Decrease	Total
Revenue	Increase	10	27	37
	Decrease	0	63	63
		10	90	100

Also noteworthy is the fact that of the one-third of plans stipulating increases in the revenue-to-GDP ratio, only about 10 were grounded in well-specified tax policy changes (such as rate increases or elimination of clearly identified tax exemptions); the rest mentioned improvements in revenue administration and tax compliance, efforts against tax evasion, or generic base-widening. Thus, less than one-sixth of our large planned consolidations envisaged deliberate increases in revenue ratios grounded in well specified measures.

This leads to an alternative perspective on large, revenue-based, fiscal adjustments that had been identified as such by the traditional strand of literature using ex-post outcomes. For example, almost all of the episodes in Alesina and Ardagna (1998) feature an increase in revenues by at least 1 percentage point of GDP; and two thirds of the observed fiscal effort in large consolidations identified by Alesina and Ardagna (2009) came from the revenue side.

How can one make sense of the different roles of revenues when contrasting ex-ante versus ex-post based definitions of large fiscal adjustments? As we will show below, in their ex-ante plan design, policymakers intended to rely on expenditure cuts rather than revenue increases (indeed, they often hoped to cut spending sufficiently to make room for tax cuts too), but the ex-post composition of adjustment turned out different than expected.

⁸ More precisely, the primary fiscal balance must be improved in response to increases in the public debt to GDP ratio (Bohn, 1998).

B. Plan Implementation

Table 2 sets out the main results on outcomes vs. plans in our sample, by looking at averages across the 66 episodes in the sample. Implementation was generally good: on average, the planned improvement in the headline (overall) balance was 2.5 percent of GDP over three-year horizons, and the actual improvement was 2.0 percent of GDP, yielding an implementation error of -0.5 percent of GDP. The implementation ratio (i.e. the ratio of actual to planned improvement in the overall balance) was 0.8. Filtering out the impact of higher-than-planned interest bill savings causes the implementation ratio (for the primary balance) to decline somewhat to 0.7, still a fairly high degree of implementation.

Table 2. Summary Statistics for Plan Implementation - By Fiscal Variable

	Variable	Δ PLAN	Δ ACTUAL	Implementation RATIO = $\frac{\Delta$ ACTUAL/ Δ PLAN (1 is perfect)	Implementation ERROR = Δ ACTUAL minus Δ PLAN (0 is perfect)
in percent of GDP	Overall balance	2.5	2.0	0.8	-0.5
	Revenues	-0.1	0.5	--	0.6
	Expenditures	-2.6	-1.5	0.6	1.0
	Primary	-2.1	-0.9	0.4	1.2
	Interest	-0.5	-0.7	1.3	-0.1
	Primary balance	2.0	1.3	0.7	-0.7
in percent of potential GDP	Cyclically-adjusted primary balance	1.6	0.9	0.5	-0.7
	Cyc. Adj. revenues	-0.4	0.5	--	0.9
	Primary expenditures	-1.9	-0.3	0.2	1.6
	<i>Memo:</i> Cyclical revenues	0.2	0.5	2.5	0.3

However, the composition of actual adjustments differs substantially from that envisaged under the plans, with policymakers not fully delivering on the promised ambitious primary expenditure cuts or the accompanying tax reductions. Specifically, on average, the plans envisaged cuts in primary spending ratios by 2.1 percent of GDP and achieved cuts amounting only to 0.9 percent of GDP. Revenues compensated in part, with an overperformance of 0.6 percent of GDP. Although the observed decline in the interest bill was largely anticipated, this item accounted for a further positive surprise of 0.1 percent of GDP. On the whole, although the plans envisaged four fifths of the adjustment from primary spending cuts and the remainder from interest bill savings, in reality less than half of the adjustment came from primary spending cuts, with revenues accounting for one quarter and the remainder from interest bill savings.

Considering implementation in cyclically-adjusted terms (the bottom panel of Table 2, with all variables expressed in percent of potential GDP), three observations stand out. First, the implementation ratio is slightly lower (0.5) than for the headline variables. (In other words, as mentioned above, cyclical conditions turned out somewhat better than expected—cyclical revenues overperformed by 0.3 percent of GDP.) Second, using cyclically-adjusted variables and potential (rather than nominal) GDP as a scaling variable further reduces the relative contribution of expenditure reductions to the overall adjustment. Although primary expenditure cuts were intended to be more than sufficient for the planned reduction in the cyclically-adjusted balance, on an ex-post basis primary expenditure cuts accounted for a

small portion of the actual adjustment, whereas increases in cyclically-adjusted revenues contributed almost two thirds of the adjustment in the cyclically-adjusted primary balance. This large contribution from cyclically adjusted revenues arose from an overperformance of almost 1 percentage point of GDP against the plans.

What factors did this unexpected overperformance of cyclically adjusted revenues stem from? Although a complete answer would require further analysis, we offer three considerations. First, on the whole, governments implemented the measures they had specified in the plans: to the extent we were able to check the implementation record of tax policy measures, few were delayed or dropped, but also few were added. (On this point, see also the case studies for the G-7 countries in Europe.) Second, efforts to widen the tax base and improve compliance seem to have had some payoff, despite the rather vague manner in which they were outlined in the plans. Third, several EU countries relied on one-off revenues, especially at critical moments in the run up to EMU (see, for example, the case studies for France and Italy). Fourth, and in our view most important, a sizable portion of the overperformance reflected non-policy related revenue increases stemming from temporary factors, such as unusually strong asset market performance, which standard cyclical adjustment methods are unable to filter out. In particular, several EU countries benefited from booming stock markets and housing price bubbles, which yielded large revenue increases prior to the global crisis that began in 2008 but, in hindsight, proved to be clearly of a temporary nature.⁹

C. Isolating Determinants of Success

We begin our exploration of the potential determinants of the degree of implementation of the plans with simple bivariate charts, to be followed in the next section by a more formal multivariate regression analysis. Specifically, we consider the possible role of the size of the planned fiscal policy change (plan ambition); the strength of fiscal rules; the starting fiscal position; surprises in economic growth and inflation; and political and electoral variables.

Plan Ambition

The plans' degree of ambition might in principle be related to the implementation ratio: one might speculate, for example, that highly ambitious plans are less likely to be implemented in full. A simple scatterplot of actual versus planned adjustment (over the plans' three-year horizon) in our sample finds no support for this hypothesis (Figure 3). The bivariate regression coefficient (excluding the five marked influential observations) is 0.8, i.e., the same as the implementation ratio reported in Table 2, indicating that the implementation ratio is no different for more ambitious plans.¹⁰

⁹ By contrast, revenue increases anchored in tangible planned tax policy changes appear to be more durable. In 8 out of the 10 cases where tax policy was discernibly tightened, more than half of the resulting increase in the revenue-to-GDP ratio endured three years beyond the end of the plan horizon.

¹⁰ Thus we do not confirm the findings put forward by previous studies according to which more ambitious plans were less likely to be implemented, using shorter sample periods at the one-year frequency. Note also that the slope coefficient in our reported regression (Figure 4) is approximately 0.8 also when we include the influential observations identified using the DFBETA procedure in Stata.

“Sticks”: Fiscal Rules

Although there is a fervent debate in the economics profession and among policymakers on the pros and cons of fiscal rules, it would seem reasonable to expect the presence of strong fiscal rules to be associated with better implementation of government plans.¹¹ In our definition of fiscal rules, we include both national rules enshrined in law (which necessitate fiscal correction once fiscal indicators approach or breach certain thresholds) and well-defined supranational rules, such as the EU’s Maastricht dissuasive and corrective arms (deficit below 3 percent of GDP and debt below 60 percent of GDP, respectively). The violation of these norms and rules is typically associated with reputational costs, but in some cases also triggers penalties.

The strength of national fiscal rules (measured by index, FRSI, compiled by the European Commission’s Economic and Financial Affairs General Directorate, DG ECFIN—see its website for a complete description) seems to translate robustly into plan implementation. Figure 4 shows the distribution of overall balance implementation ratios, both for cases where the intensity of rules (in terms of design and enforcement) was strong (shaded) and where it was weak (white).¹² Stronger rules are associated with better plan implementation.

Figure 3. Actual vs. Planned Adjustment Ambition (In percent)

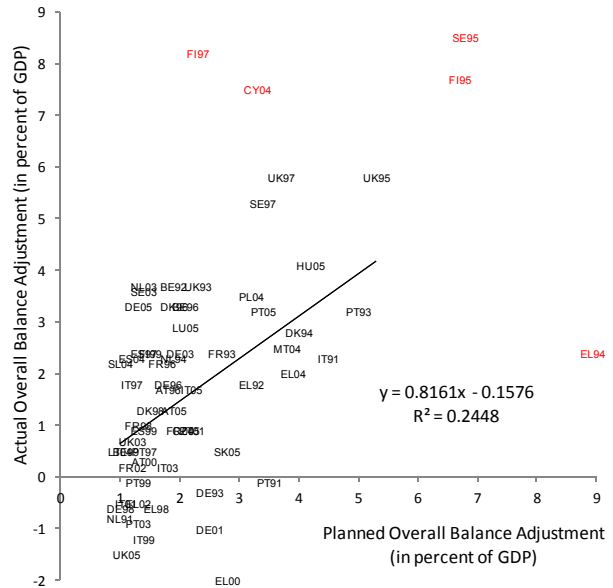
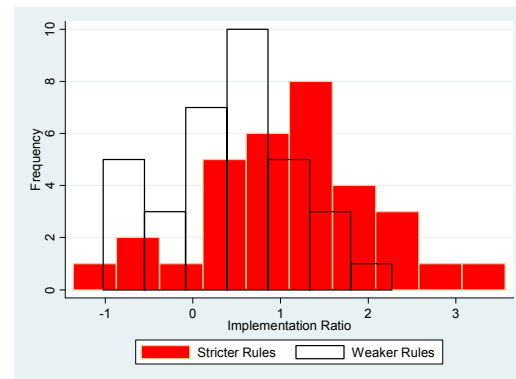


Figure 4. Does the Strength of Fiscal Rules Matter? (Stricter rules corresponds to shaded histogram)

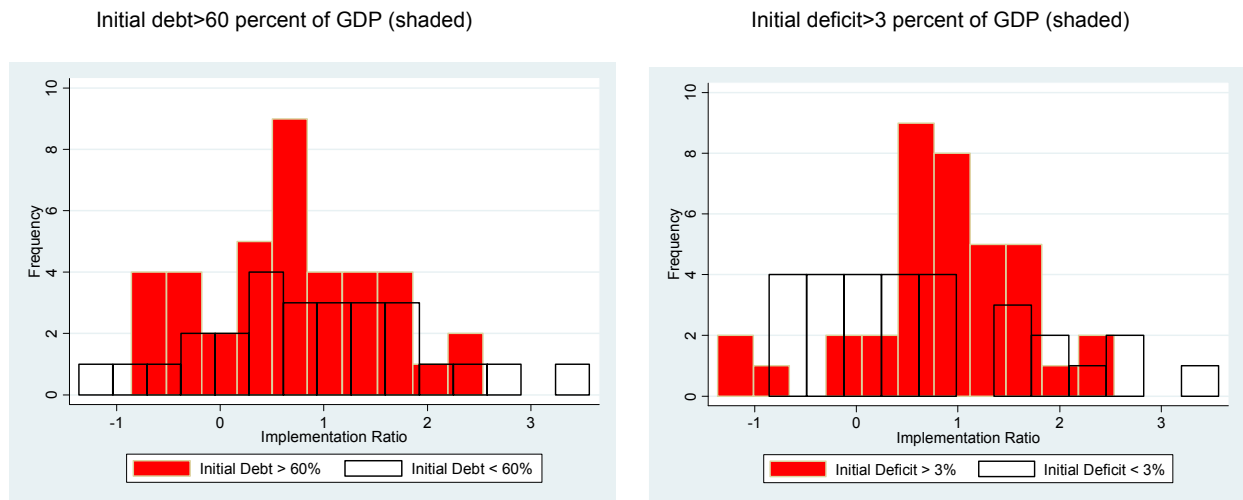


¹¹ Indeed, the proponents of fiscal rules usually argue that rules improve fiscal discipline through lower debt and deficit targets and greater compliance with such targets. Those who oppose fiscal rules point instead to the challenges involved in setting targets that appropriately reflect the economic cycle (i.e., fiscal rules lead to procyclical fiscal policy) and to increased incentives toward meeting targets through the use of accounting stratagems. Thus, many opponents of fiscal rules would probably agree that rules increase the likelihood of abiding by governmental plans. Conversely, many proponents of fiscal rules would accept that even if fiscal rules were found to improve implementation of plans, this would not settle the overall case for or against fiscal rules in their favor.

¹² Median rule strength was used as the relevant cut-off to divide the sample.

In contrast, EU-wide rules do not seem to have had much impact on the degree of implementation of plans. Indeed, Figure 5 reveals a mixed pattern at best: implementation ratios for plans prepared when initial debts were above the 60 percent of GDP limit (shaded) is not noticeably different than for plans prepared when debts were already below 60 percent. For plans prepared when deficits were above 3 percent of GDP (shaded) the frequency distribution of implementation ratios is somewhat shifted to the right compared with plans prepared when deficit ratios were already below the 3 percent limit, but this is not a statistically significant difference. This said, as shown above, the deficit limit of 3 percent of GDP may well have had an important impact on the degree of ambition of the plans, an issue we explore further later in this paper.

Figure 5. Distribution of Implementation Ratios for Maastricht Sub-Samples



Note: Non-shaded distributions refer to the plans drawn up when the initial debt was below 60 percent of GDP (left panel) and plans drawn up when the initial deficit was below 3 percent of GDP (right panel).

“Carrot”: EMU Accession

Table 3 captures the dynamics of plan implementation in the run-up to EMU accession. Because the “test date” for assessing Euro eligibility was December 31, 1997, plans formulated in 1994 were the most critical, as they had to deliver a deficit below 3 percent of GDP by end-1997.¹³ However, the 1995 and 1996 plans would also have been important, due to their proximity to the test date and the fact that EMU would only become effective at the beginning of 1999. We can clearly see the surge in planned overall balance improvement from 3 to 5 percent of GDP in 1994, and tapering off thereafter. Actual improvements followed a similar path, but always undershot planned improvements, despite stronger-than-expected growth. Even for the Euro-critical 1994 plans, only two-third of the envisaged adjustment materialized (the corresponding implementation error was -2 percent of GDP). This did not compromise the attainment of the 3 percent of GDP deficit target for 1997 because (i) the degree of ambition of the planned adjustment was set high enough to provide a “slippage cushion”; and (ii) the average 1994 deficit was revised down by more than

¹³ Appendix 2 presents a timeline for European monetary integration with key dates for fiscal policy highlighted.

1 percent of GDP in 1995 but the deficit target was maintained. Thus, the plan implementation in the run-up to EMU does not appear exceptionally strong or different, once plan ambition, growth surprises and base effects are taken into account.¹⁴

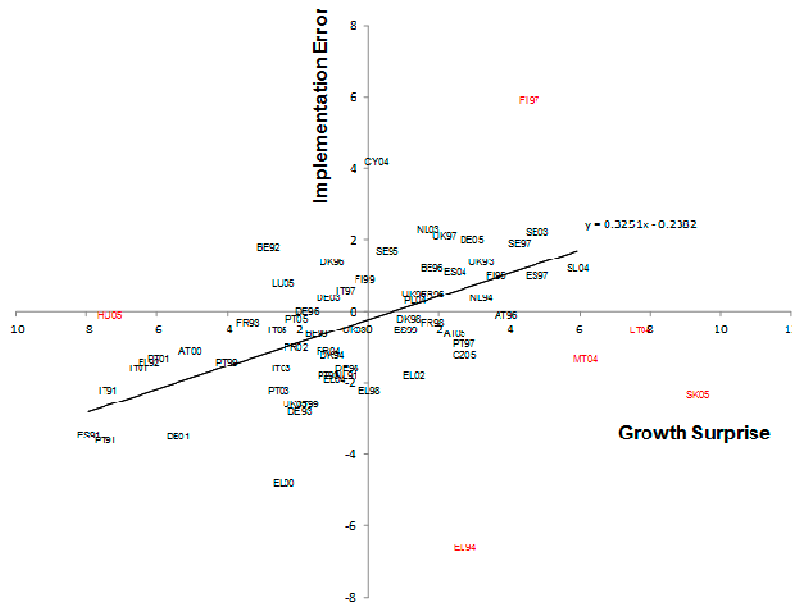
Table 3. Plan Design and Implementation in the Run-up to EMU
(Overall balance in percent of GDP; overlapping plans included; means reported)

Vintage	Plan			Actual			Actual minus Plan			
	Year <i>t</i>	Year <i>t</i> +3	Improvement	Year <i>t</i>	Year <i>t</i> +3	Improvement	Year <i>t</i>	Year <i>t</i> +3	Improvement	Growth surprise
1993	-6.0	-3.2	2.8	-5.8	-3.6	2.2	0.2	-0.4	-0.6	-0.5
1994	-7.4	-2.5	5.0	-5.5	-2.5	3.0	1.9	0.0	-2.0	1.5
1995	-4.9	-0.4	4.5	-5.6	0.0	5.6	-0.8	0.4	1.2	1.2
1996	-3.5	-0.9	2.7	-3.6	-0.2	3.4	0.0	0.7	0.7	1.8
1997	-2.2	-0.1	2.1	-2.2	1.6	3.8	0.0	1.7	1.7	2.5
1998	-1.6	-0.2	1.3	-2.1	-1.3	0.8	-0.6	-1.1	-0.6	0.8
1999	-1.0	0.4	1.4	-1.4	-1.5	-0.1	-0.4	-1.9	-1.5	-1.1

Growth and Inflation Surprises

Adverse economic growth surprises may be expected to worsen implementation ratios, through both a direct channel (weak growth implies low revenues) and an indirect channel (policymakers’ preferences shift away from fiscal adjustment and toward fiscal stimulus). Conversely, positive growth surprises, through similar reasoning, may be expected to lead to better implementation of announced plans. Figure 6 indeed suggests a strong positive association between implementation errors and growth surprises.

Figure 6. Implementation Error and Growth Surprise (In percent)



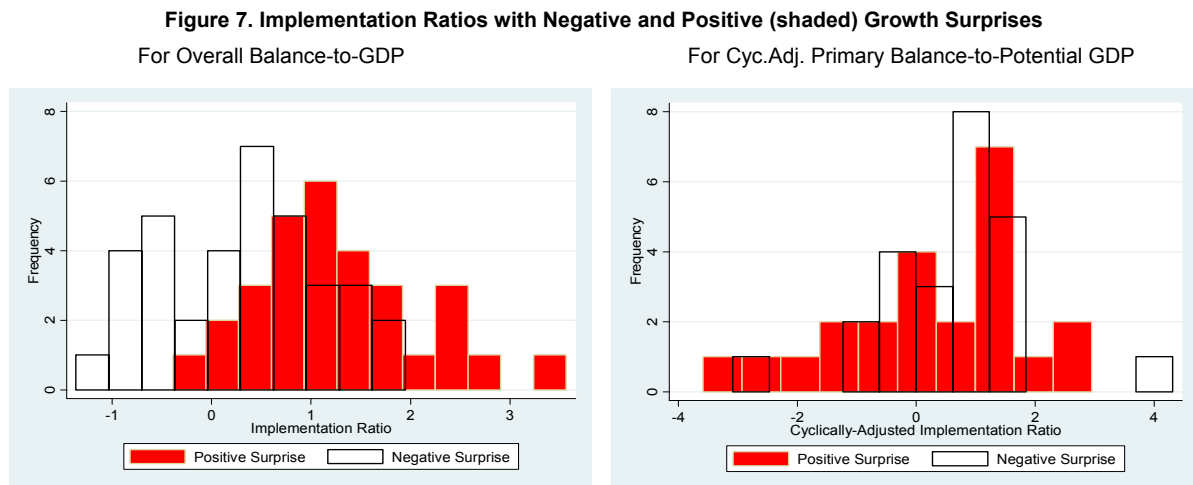
¹⁴ Testing the dummy for the euro accession or pre-97 period, we do not find any statistically significant effect.

The association is statistically significant at the conventional levels. The estimated slope coefficient, 0.3, is close to what one would expect if the only impact of growth surprises onto the implementation error were the direct impact on revenues.¹⁵

Whether the effect of growth surprises works symmetrically (i.e., with the same coefficient for both positive and negative growth surprises) is an interesting question to explore. One might expect, for example, that policymakers could plausibly spend some or all of a positive growth surprise as long as the headline deficit targets are preserved, whereas they might allow the automatic stabilizers to operate fully and even respond with stimulus measures to negative growth surprises. (In other words, policymakers might be tempted into procyclical fiscal policy in unexpected boom years but might want to use fiscal policy as a countercyclical tool against unexpected economic downturns.)

The behavioral response to growth surprises could also be cast in terms of incentives to undertake “structural” fiscal effort (i.e., improvements in the cyclically-adjusted fiscal balance) and how these are affected by growth surprises. On the one hand, an unexpected boom might afford greater political space to push through unpopular reforms; on the other hand, the incentive for fiscal adjustment through painful measures might weaken if headline balances are already improving “effortlessly” as a result of positive growth surprises. On balance, these opposing effects might well exactly offset each other.

By comparing frequency histograms for the implementation ratios associated with positive versus negative growth surprises, Figure 7 shows that growth surprises clearly have a strong impact on implementation ratios for the headline fiscal balance (left panel), but bear little relationship with implementation ratios for the cyclically adjusted primary balance (right panel). This is consistent with the view that indeed the behavioral responses to growth surprises described in the previous paragraph do approximately offset each other.



¹⁵ With revenues averaging 40 percent of GDP in our country sample and unitary elasticity of revenues with respect to GDP, the coefficient would be 0.4. A slightly smaller coefficient would reflect countercyclical fiscal measures.

At moderate inflation rates, positive inflation surprises would be expected to favorably impact fiscal adjustment via the bracket creep effect on the tax side (inflation leads taxpayers to move into higher tax brackets if these are not adjusted in a timely manner) and the real ex-post compression effect on the spending side (if spending is set in nominal terms and nominal GDP unexpectedly rises because of inflation, then spending to GDP ratios fall).¹⁶ Turning to the data, there is essentially no link between inflation surprises and plan implementation. A possible explanation for the absence of an inflation-boost for implementation could be the generally low level of inflation (around 3 percent per annum) prevailing in the European Union during the last two decades. As inflation volatility (and uncertainty) is usually associated with the level of inflation, the mean inflation surprise (in annual absolute terms) was also small during this period—a mere 0.3 percentage point, less than one-third the size of the mean absolute growth surprise. Moreover, as we work at the three-year horizon, policymakers wishing to do so had ample time to respond to inflation surprises by adjusting tax brackets and nominal spending levels.

Revisions to Initial Fiscal Position (“Base Effect”)

Large and unexpected data revisions are a fact of life for many policymakers. In the context of multi-year plans and their implementation, such data revisions, especially to the initial fiscal position, can have a significant effect on adjustment needs and incentives. For illustration purposes, consider a scenario in which the estimated fiscal balance for year t (the starting position for a three-year adjustment plan) is revised upward (i.e., unexpectedly improves) by 1 percent of GDP in year $t+1$. Policymakers could respond in several ways to this “news”: if they wish to maintain unchanged the originally-set fiscal balance target for year $t+3$, they will lower their fiscal effort in the plan’s outer years by 1 percent of GDP; but if they seek to preserve the amount of fiscal adjustment envisaged over the life of the plan, they will raise their fiscal balance target for year $t+3$ by 1 percent of GDP. In the former case, the implementation error (defined as actual minus originally-planned adjustment) would fall by 1 percent of GDP, whereas in the latter case the implementation error would not be affected by news of the revision.

The mean absolute base effect (for the year t fiscal balance-to-GDP ratio) in our sample of 66 plans is, in fact, 0.9 percentage points. This is sizable, by comparison to the average planned adjustment, for example (2.5 percent of GDP). On average, about half of the base effect showed up in plan documents within one year. This means that policymakers would have typically had sufficient basis and time to recalibrate their fiscal effort in the plan’s outer years, had they wished to do so. Although negative base effects occurred more frequently than positive ones in our sample (three-fifths vs. two-fifths), their average magnitude was similar. In principle, policymakers’ response to positive base effects could well differ from their response to negative base effects. The logic underlying this possible asymmetry could be that policymakers may be inclined to “spend” fully a positive revision to the initial deficit (thus, the implementation error would worsen by the amount of the revision), while

¹⁶ At very high inflation rates, the so-called Keynes-Oliveira-Tanzi effect would set in, whereby the real value of tax revenues falls as inflation rises. This is not relevant for the countries in our sample over the past two decades.

insufficiently compensating for negative base effects (the implementation error would remain unchanged). This hypothesis will be explored below in a formal regression framework.

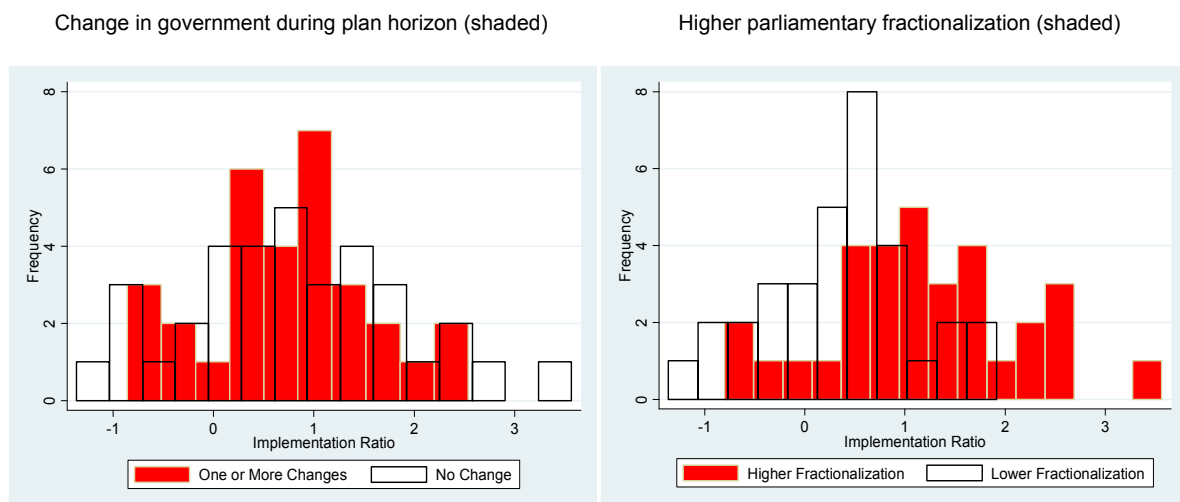
Political Variables

Several political factors might contribute to determining the plans' degree of implementation include. We focus here on two potentially important political variables for which sufficient data and intuitive hypotheses are available:

- *Change in government during the plan period*—which would generally be expected to weaken implementation, because new governments would likely be less committed to their predecessors' plans; and
- *Parliamentary fractionalization*—a high average level of parliamentary fractionalization over the plan period would likely be associated with weakened ability to implement large/painful fiscal adjustments.¹⁷ In what follows, parliamentary fractionalization is measured as the probability that two deputies randomly drawn from the legislature will belong to different parties. (The data for this variable are collected from the World Bank's Database of Political Institutions, 2010 vintage).

Figure 8 explores these hypotheses through simple charts. The left panel, which contrasts plan implementation for cases with (shaded) and without changes in government, shows no discernable difference between the two distributions. At best, one can say that changes in government do not seem to hamper plan implementation.

Figure 8. Plan Implementation Ratios for Sub-Samples of Political Variables



¹⁷ High parliamentary fractionalization would also be expected to lead to conservative growth projections. In fact, the government would seek to avoid negative growth surprises, which in turn would necessitate renegotiating the fiscal adjustment plan—a cumbersome process with a fractionalized parliament. This effect would tend to lead to positive growth surprises (von Hagen, 2010).

The right panel suggests that implementation ratios are higher for less fractionalized parliaments—a tentative relationship that we will explore further, given the need to control for other factors.

IV. REGRESSION ANALYSIS

The above descriptive analysis has shed preliminary light on possible determinants of plan implementation using simple bivariate relationships (illustrated, for the most part, through distributions of implementation “ratios” for subsets of the sample). We now turn to our more formal regression analysis, where we investigate the determinants of “implementation errors” (i.e., actual minus planned adjustment, all scaled by GDP) in a multivariate setting.¹⁸

In line with the discussion in the previous section, we explore several explanatory variables. In the first and simpler set of estimates (corresponding to Table 4), we focus on five core variables (discussed below). Additional institutional and political variables—as well as the asymmetric effects of positive and negative growth surprises (and base effects) on plan implementation—are then explored in more detailed estimates (corresponding to Tables 5 and 6).

Table 4. Baseline Regressions with Core Variable

Dependent Variable: Implementation error = actual minus planned adjustment						
VARIABLES	Fixed Effects			FE Instrumental Variables		
	(1)	(2)	(3)	(4)	(5)	(6)
Overall balance base effect	-0.61*** (0.17)	-0.61*** (0.16)	-0.53** (0.19)	-0.82*** (0.22)	-0.85*** (0.24)	-0.53*** (0.13)
Initial fiscal balance	-0.26 (0.22)		-0.38 (0.24)	-0.27* (0.16)		-0.39** (0.18)
Real GDP growth surprise	0.34*** (0.059)	0.34*** (0.059)	0.12 (0.13)	0.52*** (0.088)	0.53*** (0.094)	-0.029 (0.079)
Plan ambition	-0.37 (0.47)	-0.19 (0.32)	-0.29 (0.47)	-0.29 (0.33)	-0.14 (0.29)	-0.29 (0.33)
Deviation of initial deficit from 3% of GDP level		0.18 (0.18)			0.24 (0.19)	
Time dummies			Yes			Yes
Observations	66	66	66	66	66	66
R-squared	0.456	0.422	0.662	0.367	0.320	0.177

¹⁸ Specifically, the implementation error is defined as the difference between the actual adjustment (the difference of the *actual* overall fiscal balance-to-GDP ratio between years $t+3$ and t) and the planned adjustment (the difference of the *planned* overall balance between years $t+3$ and t). A positive implementation error, therefore, corresponds to an overperformance in adjustment relative to plan. A negative implementation error implies weaker-than-planned adjustment. As we switch from implementation ratios to implementation errors, we control for plan ambition. Using implementation errors in a regression setting has two advantages: first, marginal effects are scaled in percent of GDP, making for intuitive results; second, implementation errors are the standard dependent variable in regression-based analyses undertaken by the fiscal policy credibility literature, thus facilitating comparison with previous studies.

The five core variables are as follows:

1. *Initial fiscal balance*, measured by the year t overall balance (in percent of GDP). A larger initial balance (i.e., a better starting fiscal position) might be expected to be associated with worse implementation of fiscal adjustment plans.
2. *Fiscal balance base effect*, measured by the ex-post revision of the initial overall fiscal balance (in percent of GDP). An upward revision in the initial balance would reduce the need for adjustment effort, and could thus worsen implementation.
3. *Real GDP growth surprise*, measured by the percentage point difference between the three-year cumulative actual and planned growth rates over the plan horizon. In the absence of an active response by policymakers, the direct effect on the headline fiscal balance would be essentially equal to the automatic stabilizers (i.e., approximately the percentage point surprise in third year output times the share of total revenues in GDP).¹⁹
4. *Plan ambition*, measured by the planned improvement in the fiscal balance between year t and year $t+3$ (in percent of GDP). This explores whether more ambitious plans face challenges in implementation.
5. *Deviation of initial deficit from 3 percent of GDP* (the Maastricht/EDP threshold), measured in percentage points of GDP (and set to zero when the initial deficit ratio is below 3 percent). A larger deviation of the initial deficit from the 3 percent of GDP limit would induce greater urgency to undertake corrective fiscal action and thus be associated with stronger implementation.

The nature of our sample (unbalanced panel with 22 countries) and explanatory variables requires us to address some methodological issues. (The remainder of this paragraph is important from a technical standpoint, but can be skipped with no loss of continuity by the non-technical reader.) As in most estimations based upon panels of countries, it is likely that both the dependent variable and the explanatory variables are correlated with unobserved country characteristics. To avoid this problem, we use the standard technique of controlling for country characteristics (that do not change over time) by using the panel fixed effects (FE) estimator (with robust standard errors). Moreover, one might reasonably be concerned about two-way causation between the dependent variable (unexpected improvement in the headline fiscal balance) and the surprise in economic growth—one of the explanatory variables.²⁰ To avoid the resulting bias, we use instrumental variable estimation (FE-IV). Specifically, we use the average of other countries' real GDP growth surprises to instrument

¹⁹ For example, a 2 percentage point positive output surprise in the plan's third year in an economy where the ratio of revenues to GDP is 40 percent would result in an "automatic" improvement in the headline fiscal balance by 0.8 percentage points of GDP.

²⁰ To better understand this "endogeneity" problem, consider an actual improvement in the overall balance that manifests itself in a higher implementation error (dependent variable). Insofar as the fiscal improvement (tighter aggregate demand) exerts the expected Keynesian effects on actual output, the real GDP growth surprise (explanatory variable) would tend to be lower. This would lead to bias in the estimation.

for the growth surprise of country i (this is similar to Beetsma and others, 2009). In intuitive terms, the identifying assumption is that the only channel through which other countries' GDP growth surprises are related to the implementation error is through the GDP growth surprise in the country in question; and the implementation error in an individual country has essentially no impact on other countries' GDP growth surprise.

V. REGRESSION RESULTS

The baseline regressions in Table 1 suggest that base effects and the growth surprises are the key drivers of implementation errors (deviations of actual from planned fiscal adjustment). The initial fiscal balance seems to play a role. Other potential explanatory variables, including plan ambition or the deviation of the initial deficit ratio from 3 percent did not turn out significant.

The *base effect* turns out to be significant in all the regressions with its coefficient varying between -0.5 and -0.8. The fairly large magnitude of this coefficient would seem to suggest that policymakers do not adjust their initially-set deficit *targets* by much in the face of revisions to the initial fiscal balance, and so it is the fiscal *consolidation* which ends up adjusting relative to plan. In other words, if the initial balance is revised upward (downward) by 1 percent of GDP, actual implementation would fall (rise) by one-half to four-fifths of the revision. In some specifications, the coefficient is not statistically different from -1. In those cases, the null hypothesis that policymakers stick to the original overall balance target cannot be rejected. Note that the above interpretation implicitly assumes that the coefficient is symmetric for positive and negative base effects, an assumption that we probe further below.

For the *growth surprise*, the coefficient is about 0.3 with FE and 0.5 with FE-IV. This means that a 1 percentage point increase in the cumulative three-year growth surprise leads to an increase in the implementation error in the range of 0.3–0.5 percent of GDP. This is consistent with an average government size of about 40 percent of GDP in our sample, and an approximately unit elasticity of revenues with respect to GDP. Later, we investigate if the coefficient is similar in the case of positive and negative surprises. Importantly, the inclusion of time (individual year) dummies causes the effect of the growth surprise to disappear in both sets of regressions (Table 1, columns 3 and 6). This is to be expected, given that growth across EU countries is largely driven by common factors. Turning to the difference between the size of the coefficients across FE and FE-IV, the smaller FE coefficient is consistent with the downward bias that would arise from reverse causality, and the magnitude of the shift suggests that instrumenting for it is important.

The coefficient on *initial fiscal balance* varies between -0.25 and -0.4 and, although not consistently significant, has an intuitive interpretation: when policymakers face higher initial deficits (as seen at the time of the plan's inception) they respond with extra effort to reduce plan implementation errors. On average, this implies lower implementation errors by 0.25–0.4 percentage point of GDP for every 1 percent of GDP widening in the initial overall deficit. This estimate is similar to those reported in previous studies in the fiscal policy credibility literature.

The lack of significance of other regressors also has important implications. That *plan ambition* turns out to be an insignificant predictor for implementation suggests that

deviations of outcomes from plans are—on average, controlling for other determinants—similar for more ambitious fiscal adjustment plans as they are for less ambitious plans.²¹ Lack of significance of the coefficient on *initial deviations from the 3 percent deficit limit* indicates that EU-wide fiscal rules, while relevant for plan ambition (as shown earlier in section 3) do not substantially strengthen the plans’ implementation once they have been designed.²²

Having analyzed the core explanatory variables, we now turn to somewhat more complicated regression specifications. In Table 5, we explore specifications that allow for asymmetries in the impact of base effects and growth surprises, and the role of institutional and political variables.

Table 5. Additional Regressions with Political Variables and Asymmetries

Dependent Variable: Implementation error = actual minus planned adjustment						
VARIABLES	FE Instrumental Variables					
	(1)	(2)	(3)	(4)	(5)	(6)
Overall balance base effect	-0.82*** (0.22)		-0.83*** (0.22)	-0.84*** (0.21)	-0.68*** (0.21)	-0.87*** (0.22)
Initial fiscal balance	-0.27* (0.16)	-0.27* (0.16)	-0.31 (0.19)	-0.39** (0.16)	-0.30** (0.13)	-0.30* (0.17)
Real GDP growth surprise	0.52*** (0.088)	0.50*** (0.089)		0.49*** (0.080)	0.45*** (0.088)	0.50*** (0.087)
Plan ambition	-0.29 (0.33)	-0.26 (0.32)	-0.33 (0.33)	-0.40 (0.32)	-0.46 (0.28)	-0.29 (0.34)
Positive overall balance base effect		-0.92*** (0.34)				
Negative overall balance base effect		-0.71** (0.31)				
Positive growth surprise			0.37 (0.29)			
Negative growth surprise			0.58*** (0.16)			
Fiscal rule strength				0.61** (0.30)		
Change in government stability					3.07*** (1.12)	
Parliamentary fractionalization						-3.62* (2.09)
Observations	66	66	66	66	66	65
R-squared	0.367	0.380	0.375	0.426	0.493	0.382

Allowing for different coefficients for positive and negative base effects (column 2), we find the absolute value of the coefficient on positive (favorable) base effects to be higher than on negative ones, though the difference is not statistically significant (but will be statistically significant when observations deemed unduly influential through a standard procedure are removed—see below). On average, a 1 percent of GDP positive base effect, other things equal, worsens plan implementation by 0.9 percent of GDP, whereas a 1 percent of GDP

²¹ This finding also challenges the conclusion of earlier studies—based on annual data for a shorter sample period—that ambitious plans are associated with weak implementation (Beetsma and others, 2009).

²² As the initial fiscal balance and the deviation of the initial fiscal deficit from the 3 percent limit are strongly correlated, we avoid collinearity by never including them in the same regression.

negative base effect improves implementation by 0.7 percent of GDP. In other words, if the initial fiscal balance is found to be better than originally estimated when the plan was drawn up, the fiscal balance outcome at the end of the three years is almost unchanged; whereas if the initial fiscal balance is revised to be, say, 1 percent of GDP worse than estimated when the plan was drawn up, the deficit at the end of the three years is only 0.3 percent of GDP worse than would otherwise have been the case. These results appear to lend some support to the hypothesis that a favorable surprise about the initial fiscal position likely induces further fiscal laxity to a greater extent than an unfavorable surprise induces additional fiscal austerity.

There is also tentative (not statistically significant) evidence of asymmetry in the response to positive and negative growth surprises (column 3). The estimates suggest that policymakers essentially allow the automatic stabilizers to operate against the background of positive growth surprises (specifically, a 1 percentage point positive surprise in output by the end of the third year is associated with a 0.4 percentage point of GDP improvement in the fiscal balance). When growth surprises on the negative side, however, the implementation error increases by 0.6 percentage point of GDP, implying that policymakers undertake discretionary stimulus in addition to letting the automatic stabilizers operate fully. Note that the fiscal slippage is, in fact, greater for negative (than positive) growth surprises, which contrasts with the pattern obtained in the case of base effects (the slippage was smaller in the case of unfavorable base revisions). However, this appears intuitive, as policymakers are likely to undertake counter-cyclical fiscal measures in the face of a weaker-than-anticipated economy, whereas no such imperative or incentive exists in the case of a negative base effect.

We also find some evidence that institutional and political variables play a role. The three variables which seem to matter (are statistically significant at the conventional levels) are national fiscal rules intensity, the change in government stability and the degree of parliamentary fractionalization. Specifically, we find that a 1 point increase in the EU DG ECFIN's fiscal rules stringency index (or a one standard deviation increase that would take us from Hungary to Sweden), improves the implementation error by 0.6 percent of GDP (column 4). For changes in government stability (based upon the annual average share of veto players dropping out of government during the plan horizon, drawn from the World Bank's Database of Political Institutions, 2010 vintage) the results are also significant (column 5). If the share of veto players dropping out were increased from 0 to 20 percent, plan implementation would improve by 0.3-0.4 percent of GDP. For parliamentary fractionalization (the probability that two deputies picked at random will be of different parties—an increase indicates higher fractionalization), a one-standard deviation increase is associated with a 0.4 percent of GDP weakening in implementation (column 6).²³

We checked the robustness of our results through several exercises, including: dropping influential observations identified as such by a standard automated procedure; using overlapping plans; and using alternative definitions of “large” fiscal adjustment plans. The results are broadly similar (Table 6) running the same regressions on a sample that excludes

²³ Note that this intuitive result obtains after controlling for growth surprises which, as discussed earlier, were likely to be more positive in coalition setups because of the built-in incentive associated with projecting conservatively (avoiding the cost of renegotiating fiscal program in the event of a negative surprise).

influential observations.²⁴ As a matter of fact, the asymmetries in the impact of positive versus negative base effects (columns 2 and 6) and positive versus negative growth surprises (columns 3 and 7) become statistically significant. However, parliamentary fractionalization loses its significance.

Table 8.6 Regressions with Outliers Removed

Dependent Variable: Implementation error = actual minus planned adjustment								
VARIABLES	Fixed Effects				FE Instrumental Variables			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Overall balance base effect	-0.70*** (0.20)		-0.82*** (0.21)	-0.72*** (0.20)	-0.87*** (0.20)		-1.22*** (0.34)	-0.85*** (0.18)
Initial fiscal balance	-0.14 (0.11)	-0.13 (0.12)	-0.19 (0.11)	-0.30** (0.13)	-0.15* (0.091)	-0.14 (0.097)	-0.30** (0.12)	-0.30*** (0.10)
Real GDP growth surprise	0.38*** (0.056)	0.37*** (0.049)		0.34*** (0.048)	0.54*** (0.092)	0.52*** (0.090)		0.45*** (0.064)
Plan ambition	-0.0044 (0.19)	0.033 (0.21)	-0.076 (0.19)	-0.26 (0.18)	0.11 (0.16)	0.14 (0.16)	-0.16 (0.20)	-0.15 (0.15)
Positive overall balance base effect		-1.64*** (0.47)				-1.69*** (0.47)		
Negative overall balance base effect		-0.46** (0.18)				-0.64** (0.25)		
Positive growth surprise			0.21** (0.090)				-0.087 (0.30)	
Negative growth surprise			0.48*** (0.066)				0.82*** (0.23)	
Fiscal rule strength				0.41* (0.23)				0.38** (0.16)
Change in government stability				2.68*** (0.82)				2.25*** (0.70)
Parliamentary fractionalization				-0.47 (2.64)				-1.18 (1.67)
Observations	58	58	58	57	58	58	58	57
R-squared	0.582	0.611	0.602	0.713	0.496	0.534	0.437	0.665

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

We also used the overlapping sample of about 100 observations and found broadly similar results except that the evidence on asymmetry of coefficients was weaker. In addition, the results are not materially altered when using a cutoff of 0.5 percent of GDP (instead of 1 percent of GDP) to define large fiscal adjustment.

Finally, it is good practice to list all variables analyzed in one's research. Variables we considered beyond those on which we already reported in detail in this paper include the following: the direction/momentum of change in fiscal balances, debt, growth and real exchange rate at the time of plan inception; inflation surprises; actual and planned composition of adjustment; additional political variables (changes in government, number of years left in office, degree of polarization between the ruling party and other parties); and institutional characteristics (bureaucratic quality, corruption, law and order, poverty, and democratic accountability). For the most part, we found statistically insignificant coefficients for these variables, once the main variables discussed in the body of the paper were included in the regressions.

²⁴ We used the DFBETA procedure in STATA to identify influential observations displaying large effects on regression coefficients. The eight influential observations from a regression with the full sample of 66 were: Austria (2000), Greece (1992, 1994, 2000), Finland (1995, 1997) and the United Kingdom (1997, 2005).

VI. CONCLUSION

In this paper, we analyzed the design and implementation of large fiscal adjustment plans in the European Union during 1991–2008. We explored plan features as well as the impact of macroeconomic and fiscal conditions, and political and institutional variables, on plan implementation. Our main results are as follows.

Plans' Degree of Ambition

- Planned adjustments are more ambitious the higher the initial deficit is. However, this relationship appears to be relatively weak when initial deficit levels are low (such as below 3 percent of GDP). This is somewhat reassuring, because it is consistent with the view that, on average, policymakers seek to ensure fiscal sustainability.
- In general, implementation of plans was reasonably good and was not adversely affected by plan ambition.

Macroeconomic Factors and Revisions to Fiscal Data

- Growth surprises improve implementation of plans strongly (1% higher growth improves implementation by 0.5% of GDP). There is some evidence of asymmetry: the improvement in the fiscal balance stemming from positive growth surprises is approximately in line with the automatic increase in revenues, whereas negative growth surprises are associated with a larger worsening in the fiscal balance (compared with plans), suggesting that governments undertake stimulus measures in addition to letting the automatic stabilizers work.
- Growth forecasts are largely unbiased and close to consensus, so any growth surprises are largely unpredictable. This suggests that the main value added of fiscal councils may lie in other areas than providing an independent growth forecast.
- Inflation surprises do not matter, perhaps because inflation (and uncertainty) was low during the sample period.
- Favorable revisions to the initial overall fiscal balance seem to have little impact on the final fiscal balance (and in some estimates even lead to a worse fiscal balance than would otherwise have been the case). In the case of adverse revisions, the final fiscal balance is worse by $\frac{1}{2}$ percent of GDP for each percent of GDP by which revisions worsen the initial fiscal balance.

Composition of Fiscal Adjustment

- *Ex-ante* plans are primarily expenditure-led, but *ex-post* implementation is more balanced. Revenue increases play an important role *ex-post*, although observed overperformance does not always reflect structural effort. In fact, only one-sixth of

the plans envisaged revenue increases grounded in concrete tax policy measures. Most of these measures were implemented and resulted in revenue-to-GDP increases that were largely preserved beyond the plan horizon. Still, revenue increases reflected also one-off measures and temporary factors stemming from developments (such as asset price booms) not captured by the usual corrections for the economic cycle.

- In an interesting rejoinder with past studies on large fiscal adjustments based upon ex-post data, these findings lead us to take a more benign perspective regarding the durability of revenue-based adjustments. Indeed, when revenue-based adjustments are truly intended by policymakers and grounded in reforms, they may have better chances of being durable than one might infer from experiences that included unintended revenue increases which ultimately proved to be of a temporary nature.
- The composition of planned adjustment does not seem to drive success or failure in implementation.

Institutional and Political Factors

- EMU-related carrots and sticks have not always been effective, although there is some evidence that the run up to EMU accession drove more ambitious design and better implementation, especially for plans published in 1994–1996.
- The strength of national fiscal rules/institutions improves implementation; supranational fiscal rules (specifically the EDP) matter at the design stage, but not the implementation stage.
- Better implementation is significantly associated with less fractionalized parliaments and greater improvements in government stability, but not with changes in government.

These results have considerable policy relevance going forward. If the past is to be any guide, five observations seem to be especially relevant for the daunting fiscal adjustment that many advanced countries will have to undertake in the years ahead. First, the absence of a correlation between the degree of ambition in planned adjustment and the plans' degree of implementation suggests that policymakers need not plan conservatively simply out of a concern that more ambitious adjustments would not be implemented. Second, the major impact of economic growth surprises on implementation points to the need to spell out in the plans appropriate policy responses to possible growth surprises. Third, the large impact of revisions to initial fiscal data highlights not only the importance of high quality data and timely monitoring, but also the need to reinforce adjustment measures in the event that the starting point is found to be worse than expected. Fourth, the past challenges faced in implementing large expenditure cuts and the corresponding role of unplanned revenue increases calls for redoubling efforts to ensure that spending targets are respected and careful consideration of whether tax reforms or other lasting revenue measures would also be appropriate. Fifth, the positive results regarding national fiscal institutions—such as strength, compliance and enforcement of fiscal rules—is suggestive of the role such institution can play.

APPENDIX I: DATA ISSUES AND CYCLICAL ADJUSTMENT

A. Core Dataset

The fiscal and macroeconomic variables for plans published in 1998–2007 were drawn from European Commission’s SGP documents, compiled and provided by Beetsma and others (2009).²⁵ This dataset was extended to 1991 using 39 hard copies of convergence plans. Unlike post-98 convergence plans, the pre-98 documents are released at different times of the year. The typical plan involved an SGP convergence plan document release in June or later months of year T (defining the starting position) with a fiscal adjustment horizon spanning years T+1, T+2 and T+3. However, in certain cases, the release dates were between January and May (inclusive), so that the release year was more appropriately seen as corresponding to T+1 rather than T. In each case, the October WEO vintage of year T was used for purposes of calculating the output gap.²⁶

B. Cyclical Adjustment

The cyclically-adjusted variables, if not available in the Plan documents, were generated by assuming a revenue elasticity of 1, and an expenditure elasticity of zero with respect to the output gap.

$$\frac{R}{Y^P} = \frac{R}{Y} + \frac{R}{Y} \frac{y - y^P}{y^P}$$

total structural cyclical

where R is nominal revenues; Y and Y^P are nominal GDP and nominal potential GDP, respectively; $(y - y^P)/y^P$ is the output gap, with y and y^P denoting real GDP and real potential GDP, respectively.²⁷ Of the 66 plans comprising our main sample, 22 included information on both the output gap and structural primary balance as a share of GDP,²⁸ these variables were retrieved directly from the SGP documents. In these cases, structural revenues were computed by adding primary expenditures to the structural primary balance all in percent of GDP, thus leveraging the full information on elasticities imbedded in the Plan’s structural computations. In the remaining cases, the structural and cyclical revenues, both as a share of

²⁵ The nine variables are: headline balance, revenue, expenditure, interest payments, and debt (all in percent of GDP); nominal and real growth, and inflation (in percent); and the GDP deflator.

²⁶ A special exception was made for the United Kingdom, whose pre-1998 documents reported figures for fiscal years (ending March 31), rather than calendar years. For instance, the UK convergence plan published in March 1997 included an estimated fiscal out-turn for 1996/97 (year T) and projections for the next three fiscal years: 1997/98, 1998/99 and 1999/2000 (T+1, T+2 and T+3, respectively). When migrating to calendar year basis, the 1996/97 out-turn was booked to 1997 and the Spring 1997 WEO vintage used for structural balance computations.

²⁷ When expressed as a ratio of nominal GDP, the equivalent equation is: $\frac{R}{Y} = \frac{R}{Y} \frac{Y}{Y^P} + \frac{R}{Y} \frac{Y - Y^P}{Y^P}$

²⁸ In six additional cases, only the structural primary balance was reported.

nominal and potential GDP, were computed using output gap estimations from a Hodrick-Prescott filter.

The HP-filter smoothes a real GDP series that is constructed as follows. For each convergence plan, the entire historic series of real GDP preceding the projection year T is taken from the corresponding WEO vintage. Then, real GDP series is extrapolated over the projection period (T to T+3) by using the real growth rates for that period presented in the Plan documents, and over ten more years after T+3, by assuming the year T+3 growth rate as the long-run growth rate. A smoothing parameter of 200 is chosen to minimize the average absolute difference between the output gaps reported in the plans and the output gaps estimated using the HP-filter.

Ex-post data on structural revenues and potential GDP are taken from the European Commission's AMECO database. Any missing gaps were filled using ex-post WEO data.

	Structural Primary Balance	Output Gap	Calculation for Structural Revenue in percent of Potential GDP
Case A	Available in Plan	Available in Plan	Structural primary balance + Primary expenditure; scaling to potential GDP using "Plan" Output Gap
Case B	Available in Plan	n.a.	Structural primary balance + Primary expenditure; scaling to potential GDP using "HP-Filter" Output Gap
Case C	n.a.	n.a.	Equals Plan Revenue-to-GDP ratio; "HP-Filter" Output Gap needed to back out Cyclical Revenues in percent of potential GDP

APPENDIX II. TIMELINE FOR EUROPEAN MONETARY INTEGRATION (1990–2009)

1990	Launch of the first stage of EMU: closer economic policy coordination and the liberalization of capital movements.
1991	Agreement to the five Maastricht convergence criteria (including deficit <3% of GDP and debt <60% of GDP) to participate in EMU.
1994	Start of the second stage of EMU: creation of the European Monetary Institute (EMI). Member States required to fulfill by end-1997 the five Maastricht criteria. Plans for fiscal consolidation must show deficit <3% of GDP and debt <60% of GDP at end-1997.
1995	Madrid EU summit: The single currency is named 'the euro', and the scenario for the third stage of EMU—the introduction of the euro—is set out.
1997	The SGP agreed to at the Amsterdam EU summit, to ensure budgetary discipline among the likely EMU members. SGP gave teeth to Excessive Deficit Procedure (applying to countries with deficits exceeding 3% of GDP). A requirement for members and aspirers to submit to ECOFIN, annually, Stability programs or Convergence programs, respectively, is introduced.
May 1998	The European Council agrees to launch the third stage of EMU on Jan 1, 1999 and announces that 11 states meet the criteria to adopt the single currency: Belgium, Germany, France, Ireland, Italy, Luxembourg, the Netherlands, Austria, Portugal, and Finland. European Central Bank replaces the EMI as of 1 June 1998.
Jan 1, 1999	Start of the third stage of EMU: the euro is launched as the single currency for 11 Member States. However, the euro only exists as a virtual currency.
Jan 1, 2001	Following compliance with the Maastricht criteria, Greece becomes the 12th country to join the euro area.
Jan 1, 2002	Euro banknotes and coins are introduced in the 12 euro-area Member States.
Spring 2005	SGP revised following concerns about pro-cyclical fiscal policy in the EU. Country-specific medium-term objectives (MTOs), cast in terms of cyclically-adjusted balances, are set.
Jan 1, 2007	Slovenia becomes the 13th member of the euro area in 2007.
Jan 1, 2008	Cyprus and Malta bring the number of euro-area members to 15.
Jan 1, 2009	The euro celebrates its first 10 years, and welcomes its 16th member - Slovakia.

REFERENCES

- Alesina, Alberto and Silvia Ardagna, 1998. "Tales of Fiscal Adjustment," *Economic Policy*, Vol. 13, No. 27, pp. 489-545.
- Alesina, Alberto and Silvia Ardagna, 2009. "Large Changes in Fiscal Policy: Taxes Versus Spending," NBER Working Paper 15439.
- Baldacci, Emanuele, Benedict Clements, Sanjeev Gupta and Carlos Mulas-Granados, 2006. "The Phasing of Fiscal Adjustments: What Works in Emerging Market Economies?" *Review of Development Economics*, Blackwell Publishing, Vol. 10, No. 4, pages 612-631, November.
- Beetsma, Roel, Massimo Giuliadori and Peter Wierdsma, 2009. "Planning to Cheat: EU Fiscal Policy in Real Time," *Economic Policy*, Vol. 24, No. 60, pp. 753–804.
- Beetsma, Roel & Giuliadori, Massimo, 2008. "Fiscal Adjustment to Cyclical Developments in the OECD: An Empirical Analysis Based on Real-Time Data," CEPR Discussion Paper 6692, London, United Kingdom.
- Bohn, Henning, 1998, "The Behavior of U.S. Public Debt and Deficits," *Quarterly Journal of Economics*, Vol. 113, No. 3.
- Cimadomo, Jacopo, 2008. "Fiscal Policy in Real Time," European Central Bank Working Paper No. 919, Frankfurt, Germany.
- Giavazzi, Francesco, Tullio Jappelli, and Marco Pagano, 2000. "Searching for Non-Linear Effects of Fiscal Policy: Evidence from Industrial and Developing Countries," *European Economic Review*, Vol. 44, pp. 1259-89.
- Gupta, Sanjeev, Emanuele Baldacci, Benedict Clements, and Erwin R. Tiongson, 2005. "What Sustains Fiscal Consolidations in Emerging Market Countries?," *International Journal of Finance and Economics*, Vol. 10, pp. 307-21.
- International Monetary Fund, 2009. "Fiscal Risks: Sources, Disclosure and Management," Fiscal Affairs Department (Washington D.C.).
- Strauch, Rolf, Mark Hallerberg and Jürgen von Hagen, 2004. "Budgetary Forecasts in Europe – The Track Record of Stability and Convergence Programmes," Economic Working Papers at Centro de Estudios Andaluces E2004/42, Sevilla, Spain.
- von Hagen, Jürgen, Andrew H. Hallett, and Rolf Strauch, 2001. "Budgetary Consolidation in the EMU," *EC Economic Papers*.
- von Hagen, Jürgen, 2010. "Sticking to Fiscal Plans: The Role of Institutions," *Public Choice*, 144, pp. 487-503.