KU LEUVEN

FACULTEIT ECONOMIE EN BEDRIJFSWETENSCHAPPEN

Management, Strategy and Innovation (MSI)

The effectiveness of cyclically adjusted budget rules in the European Union

Sebastiaan Wijsman

MSI_1902

The Effectiveness of Cyclically Adjusted Budget Rules in the European Union

Sebastiaan Wijsman[1]

February 2019

Abstract

This paper presents a game-theoretical model on how revisions of the structural balance affect the implementation of the fiscal rules in the European Union (EU). The structural balance filters the nominal budget balance for influences of the economic cycle and is therefore expected to be a better indicator for fiscal discipline. However, its derivation requires assumptions and estimates on the cyclical influences and the structural balance is as a consequence revised frequently outside the governments' control. This paper assesses how this affects the effectiveness of fiscal rules. We find that the lack of control over their compliance discourages governments to set compliant budgets. Furthermore, we find that enforcers ignore the structural balance's value in their assessment of governments' fiscal discipline. They are uncertain whether noncompliance is due to governments' decisions or bad luck. As a result, undisciplined governments might be left unsanctioned, while sanctions might be imposed on disciplined governments. We assess our theoretical findings empirically using the European Commission's national fiscal rules database. However, we do not find evidence that cyclically adjusted budget rules are less effective.

Keywords: Structural balance, Stability and Growth Pact, Fiscal rules

[1] Sebastiaan Wijsman is PhD Candidate at the KU Leuven, Belgium, Faculty of Economics and Business. Email: sebastiaan.wijsman@kuleuven.be

1. Introduction

Over the years, the structural balance obtained a central role in the European Union's (EU) fiscal rules as established in the Stability and Growth Pact (SGP). While in the early years of the SGP (1998-2005), the focus on the nominal balance rule led to pro-cyclical fiscal policies, the structural balance rule was expected to solve this problem[2]. By extracting business cycle influences from the nominal balance, so-called cyclical adjustment, the structural balance is at least theoretically a better indicator for governments' fiscal performance. Accordingly, structural balance targets were established in 2005 as the so-called Medium-Term Objectives (MTOs). These country-specific objectives define both a medium-term target and a required annual improvement.

However, the use of the structural balance as fiscal indicator gave rise to a new problem. Its derivation requires assumptions and estimates on business cycle influences which makes its reliability questionable (Orphanides and van Norden, 2002)[3]. Before the fiscal year, the business cycle influences are hard to predict which complicates budgeting. Moreover, these influences remain uncertain even after the fiscal year, which leads to frequent ex-post revisions of the structural balance estimations. These revisions are based on updated expectations about the economy's potential or stance in the cycle. They change the structural balance, outside the government's control. In section 2 we will show that the revisions are substantial compared to the reference values of the structural balance rule.

Accordingly, scholars argued that the structural balance made the SGP too complex and hinders its working. For instance, incorrect estimates can lead to inaccurate policy recommendations (Claeys, Darvas, and Leandro, 2016) and decrease the incentive for governments to comply with the fiscal rules, since institutions and the public cannot hold governments accountable for noncompliance (Cordes et al., 2015). These effects are suggested by the literature, but an in-depth assessment of the impact on the SGP's effectiveness remained absent. Therefore, this paper focuses on the incentives politicians have under fiscal rules with an unreliable indicator. How does an unreliable indicator affect government compliance with fiscal rules? And how is the enforcement of the fiscal rules affected?

The contribution of this paper is threefold. First, we assess how much uncertainty the structural balance as fiscal indicator creates. Therefore, we compare the cyclical components estimated prior and after the fiscal year. The cyclical component is the part of the nominal budget balance that is corrected to account for business cycle influences and to derive the structural balance. We look to these estimates for the fiscal years 2005-2016 for the nineteen euro area member states. While these member states are required

[2] The structural balance obtained a central position through the Commission Communication on "Strengthening the co-ordination of budgetary policies" of November 2002. The Council of the EU approved this in its ECOFIN Council Report 6877/03. The SGP's 2005 Reform included the structural balance rule formally.
[3] See Cimadomo (2014) for a literature review.

to improve their structural balance with 0.5% of GDP in normal economic times[4], the cross-country average up- or downward revision of the cyclical component ranges between 0.2% and 0.5% of GDP. During the turbulent crisis years 2008 and 2009, the cross-country average revision was even 1.1% and 1.2% respectively. Since the revisions directly influence the structural balance outside the governments' control, we conclude that the structural balance is a highly uncertain fiscal indicator, even in normal economic times.

Second, we present a game-theoretical model that assesses how governments and fiscal rule enforcers take this uncertainty into account. Our model shows that the uncertainty undermines the fiscal rule's effectiveness. In particular, we distinguish compliance and enforcement effects. First, given that governments differ regarding their fiscal discipline, we find that the uncertainty decreases the number of governments that would limit their deficits in order to avoid sanctions. Second, we find that the fiscal rule's enforcer will ignore the government's (non)compliance in its assessment of the government's fiscal discipline. This leads to inaccurate enforcement of the rules: sanctions might be imposed on disciplined governments while undisciplined governments might be left unsanctioned. This makes the structural balance's unreliability favorable for the undisciplined and unfavorable for disciplined governments.

Third, we assess the theoretical findings empirically. Using the European Commission's Fiscal Rules Database, we test how cyclical adjustment affects the effectiveness of fiscal rules. Our database consists of EU-28 between 1997 and 2016. However, we do not find evidence that cyclically adjusted budget rules are less effective. We contribute to existing empirical work in two respects. We use a year and country-specific index to measure the extent in which member states use cyclically adjusted budget rules. By contrast, Debrun et al. (2008) and Ayuso-i-Casals et al. (2009) use the index to distinguish two groups of member states losing variance over time. We were also able to enhance the number of observations to 525.

This paper is structured as follows. In the next section, we explain how the structural balance is embedded in the SGP and assess by means of cyclical component estimates how much uncertainty the indicator creates. In section 3 we provide a literature review on the impact of the uncertainty on the implementation of the fiscal rules, in section 4 we present a game-theoretical model to assess the uncertainty's effect on the working of fiscal rules, and in section 5 we present our empirical test. Section 6 provides the empirical results and we conclude in section 7.

[4] In section 2 we specify the reference values that apply during other economic conditions.

2. The structural balance and its revisions under the SGP

The structural balance is established in the Preventive Arm of the SGP to monitor governments' fiscal performance over the cycle. The initial SGP as introduced in 1998 required governments to achieve a medium-term budget position "close-to-balance or in surplus", but this objective was not explicitly defined by the Pact. In the early years of the SGP, governments maintained deficits despite the EU average real GDP growth of 3.1% between 1998 and 2004. Both the Council and the Commission therefore aimed to strengthen the Preventive Arm to ensure fiscal consolidation in periods of economic recovery[5]. The 2005 reform introduced the structural balance as fiscal indicator operationalizing the medium-term budget objective accordingly.

The structural balance targets, referred to as Medium-Term Objectives (MTOs), are country-specific objectives to be achieved over the medium term. Member states establish their own MTOs in their Stability or Convergence Programmes[6] which they submit every spring in the context of the European Semester. However, the SGP lays down country-specific minimum MTOs based on debt levels and potential expenditures related to ageing populations. MTOs must be revised every three years and after major structural reforms.

In addition to the MTO, the Preventive Arm proscribes a minimum adjustment path towards the MTO. This adjustment path refers to a required minimum annual improvement of the structural balance. In normal economic times the required annual improvement is 0.5% of GDP, but depending on the economic conditions and debt levels other thresholds may apply. In economic bad times it decreases to 0.25% or 0% depending on a member state's debt level. In economic good times, minimum improvements of 0.75% or 1% may apply. The Vade Mecum on the SGP provides an overview of the required improvements[7].

When a government deviates from its MTO or adjustment path, the Commission may initiate a significant deviation procedure. This gives the government an early warning that it may end up in an Excessive Deficit Procedure if the deviation is not corrected. Since the 2011 reform of the SGP, the Commission may propose that a member state showing a significant deviation from its adjustment path makes an interest bearing deposit of maximum 0.2% of GDP. The Council adopts this sanction by reversed qualified majority or amends by qualified majority.

[5] See the Council "Declaration on the Stability and Growth Pact" on 18 June 2004 and the Commission Communication "Strengthening economic governance and clarifying the implementation of the Stability and Growth Pact" on 3 September 2004

[6] Euro area member states submit Stability Programmes and non-Euro area member states submit Convergence Programmes.

[7] See European Commission (2018)

Despite the 2005 reform formalized the rules under the Preventive Arm, it is often interpreted as weakening of the Pact. To enhance ownership, member states are to set their own MTOs, but this also provides discretion. Moreover, the reform added escape clauses to the Corrective and Preventive Arm enhancing flexibility and making the Pact less transparent (Morris, Ongena, and Schuknecht, 2006). For instance, member states are allowed to deviate temporary from the MTO and the adjustment path towards the MTO when structural reforms are implemented.

2.1 The uncertainty of the structural balance

The SGP requires member states to adjust their structural balance towards their MTOs with a minimum annual improvement. However, the structural balance has been criticized for the uncertainty it creates. To calculate the structural balance one needs among others the output gap: The economy's stance in the cycle. This variable is unobservable and requires assumptions and estimations. In the following, we show how much uncertainty the structural balance creates. We focus on the accuracy of structural balance estimates by assessing the revisions of the cyclical component. The cyclical component is the part of the nominal budget balance that is subtracted to account for the cyclical influences on the budget. We compare the size of these revisions with the structural balance targets as established in the SGP. We find that the revisions are substantial and enhance considerable uncertainty.

For the understanding of our analysis, we first show how the structural balance is derived. Figure 1 shows that the nominal budget balance is decreased with the so-called cyclical component and one-off and temporary measures. The Commission provides some examples of one-off and temporary measures in its SGP's Code of Conduct including tax amnesties, sales of non-financial assets, or emergency costs from natural disasters[8]. However, our main focus here is the cyclical component.

Figure 1 Derivation of the structural balance

Nominal budget balance
(-) Cyclical component
= Cyclically adjusted balance
(-) One-offs and temporary measures
= Structural balance

Subtracting the cyclical component excludes the part of the balance which would change over time according the business cycle. During economic slowdowns, governments incur costs resulting from low GDP growth including higher social expenditures and lower tax revenues. The cyclical component

[8] See European Commission (2016)

corrects the nominal balance for these temporary influences. Its value can be either positive or negative and is higher during economic slowdowns and lower during booms.

However, the cyclical component is unobservable making the calculation of the structural balance uncertain. Two predictions are needed to compute the cyclical component according the official EU methodology: (*i*) the output gap (i.e. where the economy stands in the cycle) and (*ii*) how budget components react to the cycle[9]. Calculation of output gaps requires assumptions on potential GDP and also the sensitivity of budget components is not directly observable[10].

Due to the constant flow of new information and renewed assumptions the cyclical component is frequently revised, even after the fiscal year. These revisions directly translate into changes of the structural balance outside the government's control. More specifically, the Commission revises the cyclical components twice per year: for the spring and autumn editions of its European Economic Forecast (EEF). The European Commission's fiscal surveillance is accordingly based on the most recent updates. For instance, its assessment of the member states' Convergence and Stability Programmes in May is based on the EEF spring edition, whereas the assessment of the Draft Budgetary Plans in November is based on the EEF autumn edition.

2.2 Cyclical component revisions

To measure the uncertainty enhanced by the cyclical adjustment, we assess the extent to which the Commission revised its estimations of the cyclical component. We compare the annual cyclical components before and after the fiscal years. We calculate the difference between these estimates for the period 2005-2016 for the nineteen euro area members. For the cyclical components estimated prior the fiscal year t, labeled cc_{t-1}, we use the autumn edition of the Commission's EEF. These autumn estimates match our research interest since their timing coincides with the government budget preparation. For the estimates after the fiscal year, labeled cc_{t+1}, we use the spring edition of the EEF. For example, for the revision of the cyclical component over 2008 we compare the Commission's autumn 2007 with the spring 2009 estimate.

We calculate the individual revisions as $r_t = cc_{t+1} - cc_{t-1}$. This means that when r_t is positive (an upward revision), the economy was worse than the Commission expected, whereas a negative r_t (a downward revision) indicates an initial underestimation of the economy. Since we compute annual cross-country averages of the revisions, upward and downward revisions might cancel each other out,

[9] This corresponds to the "two-step methodology" of cyclical adjustment. This approach is used by most international organizations including the Commission, OECD, and IMF (Mourre et al., 2013)
[10] This parameter is periodically estimated by the OECD and agreed by the Output Gap Working Group of the Economic Policy Committee (Mourre et al., 2013)

however. Therefore, we also express the revisions in absolute terms. Our data ranges from 2005 to 2016 and we have nineteen countries in our dataset and 221 observations[11]. Figure 2 provides an overview on the distribution of the individual revisions.[12]

Figure 3 shows the average revision per year. Since most years contain both upward and downward revisions (positive and negative figures) the numerical value of the averages is flawed, but the figure provides information on the general direction of the revisions. For the years 2005-2007, the Commission revised the cyclical components slightly upwards indicating a small overestimation of the economy. However, with an average upward revision of 1.11%, the overestimation for 2008 was large. These large revisions over 2008 can be attributed to the financial crisis. The Commission did not foresee this event which affected both GDP growth in 2008 and the expected potential GDP. Over 2009, we see a large downward revision of 1.17% indicating that the Commission was too pessimistic in its autumn 2008 forecast. Between 2010 and 2016, the average revisions were smaller.

Figure 4 shows the average revisions in absolute terms. This does not provide information on the direction of the revisions, upward or downward, but it shows the general unreliability of the estimates. Comparing the absolute revision averages over the years, we see again the 2008 and 2009 peaks, but it is more interesting to look at the absolute revisions in the more predictable years. We see that even in these years, 2005-2007 and 2010-2016, the Commission revised the cyclical components on average for about 0.2% (2016), 0.3% (2005, 2006, 2012, and 2014), 0.4% (2011, 2013, and 2015), and 0.5% (2007 and 2010). For comparison, the dotted line represents the compliance threshold (0.5%) for the structural balance improvement under the Preventive Arm. The average revision during crisis years exceeds this threshold and also the non-crisis revisions are close to the actual threshold.

Finally, we conduct t-tests to assess whether the ex-post cyclical component estimates differ significantly from their ex-ante counterparts. The null hypothesis is that the absolute revision is zero. To ensure that the result is not driven by crisis years, we conduct t-tests for both the full sample and by year. Table 1 shows the results. The results indicate clearly that the revisions are high. In both crisis and non-crisis years the ex-post estimate significantly differs from the ex-ante estimate.

[11] The 2005 data for Estonia, Slovakia, Slovenia, Malta, Cyprus, Latvia, and Lithuania is missing. The Commission did not provide estimates for these countries in the European Economic Forecast Autumn 2004.

[12] The calculation of cyclical components is based on two elements: the output gap and the elasticity of budgets regarding changes in the output gap. Mourre, Astarita, and Princen (2014) show that revisions in the elasticity of budgets only has a marginal impact on the cyclical component. Using their table 4.1, we find that this impact is 0.05% on average for our set of countries. Since most of the cyclical component revisions are much higher than 0.05%, we can attribute revisions almost fully to the updated output gap estimates, which are outside the governments' control.

In short, we find that the cyclical component revisions are substantial. To put the order in perspective, consider the implications for the structural balance. The revisions translate directly into changes in the structural balance: when the cyclical component is revised upwards by 0.2%, the structural balance is decreased with 0.2%. The structural balance thus fluctuates based on new expectations about the economy's stance. While member states are in normal economic times required to obtain an annual improvement of 0.5%, the cross-country average revision in non-crisis years ranges from 0.2% to 0.5%. Since these revisions are outside the governments' control, the main lesson from our assessment is that the structural balance creates substantial uncertainty. In our model in section 4 we show the uncertainty's consequences for the working of fiscal rules.

Figure 2 Distribution of the Commission's cyclical component revisions as % of GDP

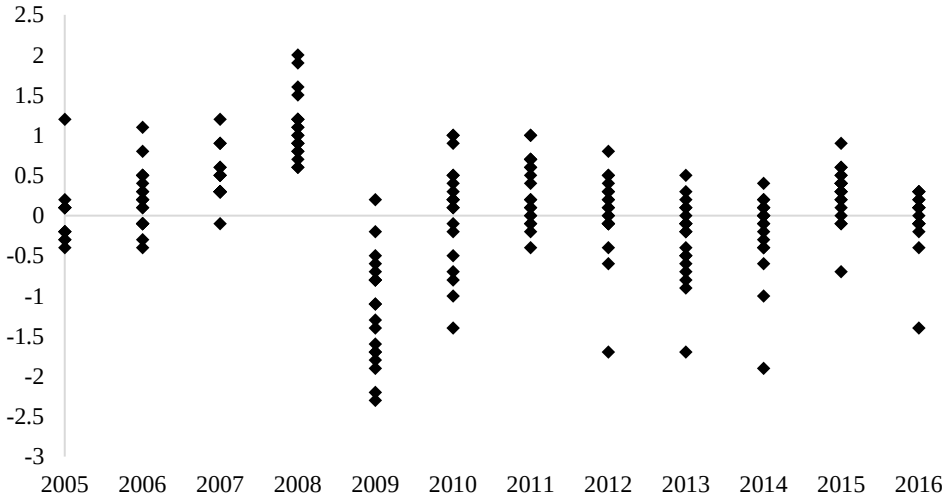


Figure 3 The Commission's average cyclical component revision as % of GDP

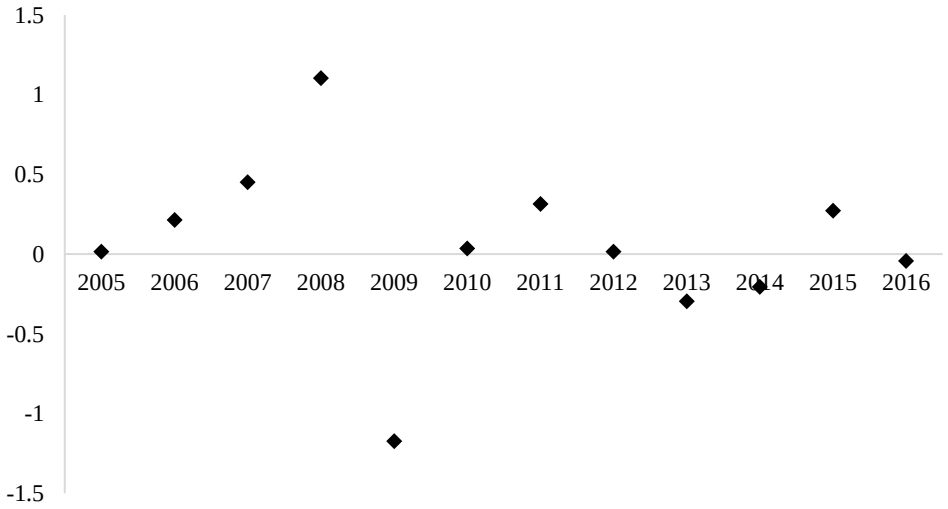


Figure 4 The Commission's average cyclical component revision as % of GDP (absolute terms)

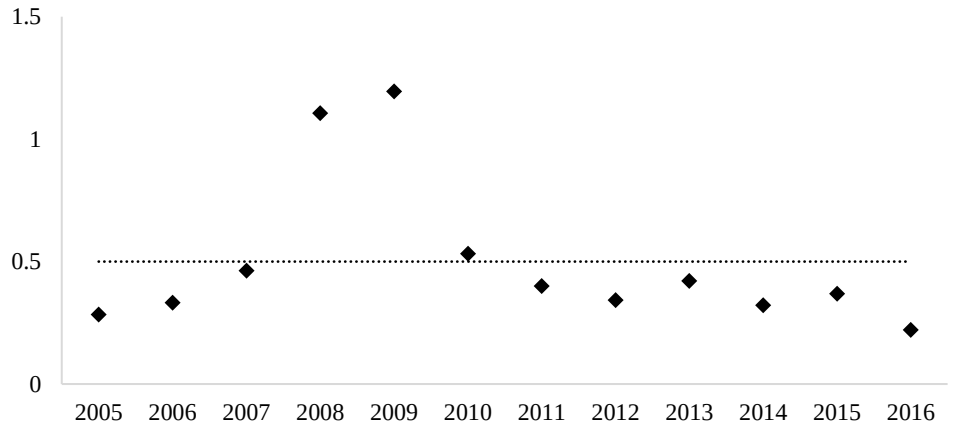


Table 1 Results t-test of cyclical component revisions (absolute terms)

T-test hypothesis: the mean of the absolute revisions is zero			
Sample	Mean	T-value	Observations
Full sample	0.51	15.58***	221
2005	0.28	3.23***	12
2006	0.33	5.42***	19
2007	0.46	7.33***	19
2008	1.11	12.07***	19
2009	1.19	8.14***	19
2010	0.53	5.96***	19
2011	0.40	5.40***	19
2012	0.34	3.78***	19
2013	0.42	4.50***	19
2014	0.32	3.07***	19
2015	0.37	6.95***	19
2016	0.22	3.16***	19

Note: *, **, *** denote significance at the 10, 5, and 1% level respectively.

3. Literature review

The aim of the structural balance is to eliminate business cycle influences from the budget balance and is therefore undoubtedly a good concept in theory. It is an old understanding that fiscal developments are affected by both permanent influences and economic fluctuations. Already in the 1930s economists developed fiscal indicators comparable to the structural balance.[13] Nowadays, the structural balance is widely used in the empirical literature to measure governments' fiscal performance or discretionary

[13] See Larch and Turrini (2010) or Costantini (2015) for the historical development of the cyclically adjusted balance.

changes in fiscal policy (see e.g Alesina and Perotti, 1995; Alesina and Ardagna, 2010). In the context of the SGP, the structural balance allows the Commission to assess fiscal discipline in any stage of the business cycle. Eyraud and Wu (2015) show that despite its measurement error the structural balance is still 25 percent more accurate in measuring governments' fiscal positions than the nominal balance.

Over time, however, research revealed the shortcomings of the structural balance and nowadays the indicator is criticized heavily for the uncertainty we showed above (see e.g. Larch and Turrini, 2010; Hughes Hallett, Kattai, and Lewis, 2012). The structural balance is more and more considered unfit for its role as key indicator in the SGP's Preventive Arm (Claeys, Darvas, and Leandro, 2016). The literature argues that the structural balance decreases the enforceability of the SGP and makes it overly complex. The estimations and assumptions give room for disagreement, lead to conflicts over technicalities, and complicate the compliance monitoring (Schuknecht, 2004; Schaechter et al., 2012).

Before the fiscal year, governments are uncertain whether draft budgets will match the requirements related to the structural balance, since the output gap and cyclical component are still unknown (Cimadomo, 2012). After the fiscal year, the Commission is uncertain whether structural balances are (non)compliant due to governments' fiscal (in)discipline or due to changes in the economic cycle (Cordes et al., 2015). Governments may use this uncertainty in an opportunistic manner and leverage it to undermine enforcement (Hughes Hallett, Kattai, and Lewis, 2012).

The uncertainty is in particular detrimental since the SGP relies to a large extent on soft law (Schuknecht, 2004; Buti, 2016; Schaechter et al., 2012). There are formal enforcement tools, but peer pressure from other member state governments also works disciplining. However, the room for discussion as enhanced by the uncertainty affects the SGP's political sphere, reducing peer pressure (Schuknecht, 2004). Moreover, the technicalities also negatively affect the pressure voters put on their governments to comply with the rules. The public debate does not focus on technicalities related to cyclical adjustment but only notices issues that are easily understandable (Schuknecht, 2004). Finally, Commission recommendations for fiscal policy adjustment can be misguided ex-post (Wolff, 2017; Bénassy-Quéré et al., 2018).

3.1 Empirical evidence

We would thus expect that the uncertainty affects the implementation of fiscal rules negatively and some papers find related evidence. Focusing on expenditure rules, Cordes et al. (2015) find that governments are more inclined to comply with rules based on indicators that are more under their control. They find that specific expenditure ceilings have the highest compliance rate while expenditure rules specified in terms of GDP, change to GDP, and real expenditure growth have lower compliance rates. Governments

do not fully control GDP and inflation which makes expenditure levels denoted in percentage of GDP or real growth uncertain. This finding suggests that the uncertainty enhanced by the structural balance makes governments less inclined to comply with the structural balance rule.

Debrun et al. (2008) and Ayuso-i-Casals (2009) find that cyclically adjusted fiscal rules have less impact on fiscal discipline than other rules. Both papers assess for EU-25 the impact of several fiscal rule design features on fiscal performance between 1990 and 2005. By means of a Fiscal Rule Cyclicality Index they found that fiscal rules that take business cycle influences into account have a weaker disciplining effect than rules that do not. Our empirical analysis in section 5 builds on these studies.

Based on this literature, we expect that structural balance revisions affect both fiscal rule enforcers and governments. Enforcers will take the revisions into account when deciding whether to sanction noncompliance. We expect that therefore the enforcer may make enforcement mistakes as it cannot rely on the information provided by the structural balance. On the other hand, governments may exploit the uncertain nature of the structural balance to their advantage. In particular, we expect that governments are not encouraged to pursue higher budget balances when the structural balance may not reflect their fiscal (in)discipline. In sum, we expect that the structural balance's uncertainty weakens the disciplining effect of fiscal rules. In our game-theoretical model, we assess the effect in depth.

4. The model

In this section we present a game-theoretical model on the role of fiscal indicators in compliance and enforcement of fiscal rules. The aim of this model is to assess the impact of the structural balance revisions on the enforcement of the fiscal rules and the corresponding government behavior being aware of the revisions. In the model there is a government setting a deficit and an enforcing actor that may sanction excessive deficits. The enforcer cannot observe the deficit directly, but may rely on the information provided by a message about the deficit mimicking a fiscal indicator like the structural balance.

4.1 The setup

The game consists of one period. In this period, a government must decide on its deficit b and an enforcer must decide on whether to impose sanctions on the government. In line with the European context, we refer to the enforcer as 'the Commission'. However, this could have been any actor that enforces fiscal discipline like financial markets or domestic voters (Cordes et al., 2015).

The government and the Commission have preferences over the deficit level. The preferred deficit of the government depends on its type j. A disciplined government, denoted by $j = d$, prefers deficit b_1 and its utility function is accordingly:

$$V_d = -(b_1 - b)^2 - C$$

On the contrary, an undisciplined government, denoted by $j = u$, prefers deficit b_2:

$$V_u = -(b_2 - b)^2 - C$$

We assume that $b_2 > b_1$ and, for simplicity, the government can only set b_1 or b_2[14]. C represents the sanction that may be imposed by the Commission. C is strictly positive if the Commission decides to sanction and zero otherwise. The sanction may represent all sorts of repercussions like pecuniary sanctions, official warnings, or the loss in reputation. The government type is exogenously determined. The probability that there is a disciplined government in office is q, and the probability that there is an undisciplined government in office is $1 - q$.

To promote fiscal discipline the Commission may sanction the government after the deficit is set. Like the disciplined government, the Commission prefers deficit level b_1[15]. Its utility function is accordingly:

$$W = -(b_1 - b)^2 - S - R$$

S represents the enforcement costs which is strictly positive if the Commission sanctions a disciplined government and zero otherwise. R represents the Commission's reputation costs which is strictly positive if the Commission leaves an undisciplined government unsanctioned and zero otherwise. A possible interpretation of these costs are the political consequences for the Commission. National populations will not appreciate the Commission sanctioning disciplined governments or leaving undisciplined governments unsanctioned.

When deciding on sanctions, the Commission cannot directly observe the deficit level. To inform the Commission and to facilitate its sanctioning instead, there is a message sent by nature providing information on the government's deficit. This message serves as fiscal indicator. When the government set the lower deficit b_1, message $m = b_1$ is sent to the Commission, whereas $m = b_2$ is sent when the government set the higher deficit b_2. However, in line with the structural balance's revisions as shown in section 2, the message is sent with an error. The message $m = b_2$ is sometimes sent when the

[14] We refer to deficit levels and for simplicity we ignore the underlying government income and expenditures. For the model it is essential that there are two government types of which one's deficit preferences are not aligned with the enforcer's.

[15] The notion that one government type has the same preferences as the monitoring actor is also used in Chaudoin (2014). There, the 'good' and 'bad' types differ in terms of their preferences over trade barriers.

government set deficit b_1 (a false positive) while sometimes the message $m = b_1$ is sent while the government set deficit b_2 (a false negative). This distinction of error types is introduced in the SGP's context by Hughes Hallett, Kattai, and Lewis (2012) and mimics the inaccuracy of structural balance estimates. The size of the error is denoted by $\alpha > 0$ and will capture both error types. The incorrect message is sent with probability α and the correct message with probability $1 - \alpha$. Accordingly, the probability distribution of the message for each deficit level is summarized below:

$$P(m = b_1) = \begin{cases} 1 - \alpha & \text{if } b = b_1 \\ \alpha & \text{if } b = b_2 \end{cases}$$

$$P(m = b_2) = \begin{cases} \alpha & \text{if } b = b_1 \\ 1 - \alpha & \text{if } b = b_2 \end{cases}$$

We assume $\alpha < \frac{1}{2}$ to exclude the situation that there is a lower probability that $m = b_2$ is sent under the higher deficit. Otherwise, disciplined governments would pursue high deficits to minimize the probability that the higher deficit is messaged, which is unrealistic.

There is asymmetric information in this model. The government knows its type and the preferred deficit of the Commission. The Commission is aware of the deficit levels preferred by the disciplined and undisciplined government, but does not know which government type is in office. Moreover, the Commission cannot infer the deficit before the end of the game. Both the government and the Commission can observe the message and are aware of the error. Also C, S, and R are known by all actors.

The sequence of events is depicted in figure 5 and the game tree in figure 6. First, government type j is exogenously determined. The government observes its type, and sets deficit b. Then, nature sends the message m according the probabilities. The Commission observes the message, and determines whether to impose sanctions C or not. Finally, the game is ended and pay-offs are allocated.

Figure 5 Sequence of events

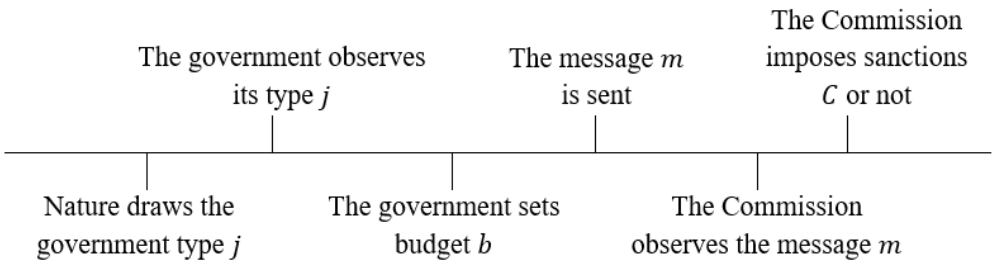
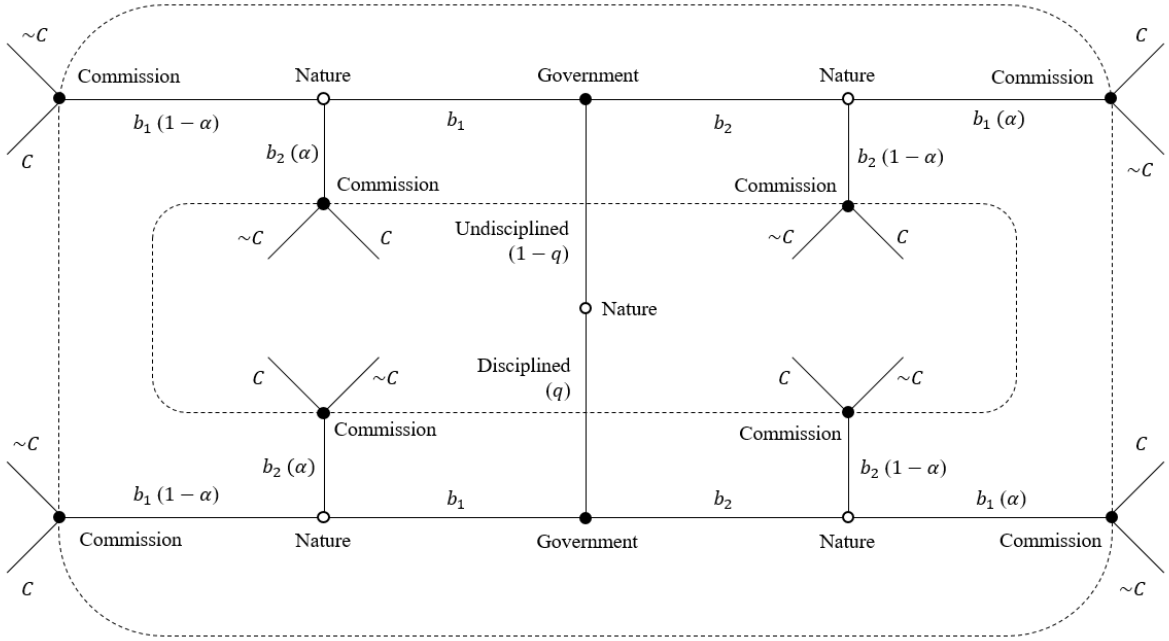


Figure 6 The game tree (with probabilities between parentheses)



4.2 The equilibrium

We are looking for a perfect Bayesian equilibrium. The strategy profile of the government, regardless of its type, consists of a deficit b, whereas the Commission's strategy profile comprehends a sanction strategy that establishes under which conditions it will impose sanctions C. The equilibrium strategies must be consistent with the actors' beliefs over the state of the world and we use backward induction to find the perfect Bayesian equilibrium.

First, suppose that both government types always set their preferred deficit levels. Accordingly, the disciplined government sets deficit $b = b_1$ and the undisciplined government $b = b_2$. Consequently, under the disciplined government the correct message $m = b_1$ is sent with probability $1 - \alpha$ and the incorrect message $m = b_2$ with probability α. Under the undisciplined government the correct message $m = b_2$ is sent with probability $1 - \alpha$ and the incorrect message $m = b_1$ with probability α.

Furthermore, we propose the Commission's intended strategy to sanction the undisciplined type, but not the disciplined type. This is a rational intention given the Commission's enforcement costs S if it sanctions the disciplined type and reputation costs R if it leaves the undisciplined type unsanctioned. However, based on the governments' deficit strategies and the message errors, the Commission cannot infer the government's type with certainty. When the message indicates that the higher deficit is set ($m = b_2$), the government in office might be disciplined or undisciplined. Therefore, the Commission

compares its expected utility from sanctioning with its expected utility from not sanctioning. When we let E indicate that we refer to the expected utility, the Commission will impose sanctions if:

$$(1) \qquad EW_c > EW_{\sim c} \quad \Rightarrow \quad P(j = u|m = b_2)R > P(j = d|m = b_2)S$$

The Commission uses the message and Bayes' rule to update its beliefs over the government's type. After receiving the message, the Commission updates its beliefs into:

$$P(j = d|m = b_2) = \frac{\alpha q}{\alpha q + (1-\alpha)(1-q)} \qquad P(j = d|m = b_1) = \frac{(1-\alpha)q}{(1-\alpha)q + \alpha(1-q)}$$

$$P(j = u|m = b_2) = \frac{(1-\alpha)(1-q)}{\alpha q + (1-\alpha)(1-q)} \qquad P(j = u|m = b_1) = \frac{\alpha(1-q)}{(1-\alpha)q + \alpha(1-q)}$$

When we insert these probabilities in the sanctioning condition (1), we find that the Commission will impose sanctions after receiving the message $m = b_2$ if:

$$(2) \qquad S < R\frac{(1-\alpha)(1-q)}{\alpha q}$$

and after receiving the message $m = b_1$ if:

$$(3) \qquad S < R\frac{\alpha(1-q)}{q(1-\alpha)}$$

Note that (2) always holds when (3) holds, because $\alpha < \frac{1}{2}$. Accordingly, we distinguish three cases. First, the Commission always sanctions, that is when (3) holds. Second, the Commission never sanctions, that is when neither (2) nor (3) holds. Third, the Commission only sanctions after receiving the message $m = b_2$, that is when (2) holds but (3) does not. Table 2 depicts the (expected) utilities following from the set of proposed strategies.

Table 2 Expected utilities from the proposed strategies

(2) and (3) hold (The Commission always sanctions)	$V_d = -C$ $V_u = -C$ $EW = -Sq - (1-q)(b_1 - b_2)^2$
(2) and (3) fail (The Commission never sanctions)	$V_d = 0$ $V_u = 0$ $EW = -(1-q)(b_1 - b_2)^2 - (1-q)R$
(2) holds and (3) fails (The Commission only sanctions after receiving message $m = b_2$)	$EV_d = -\alpha C$ $EV_u = -(1-\alpha)C$ $EW = -q\alpha S - (1-q)(b_1 - b_2)^2 - (1-q)\alpha R$

Note: E indicates that we refer to an expected utility.

In Appendix A we show that this is an equilibrium for the Commission and the disciplined government. The undisciplined government, however, might have incentive to deviate from its proposed strategy if (2) holds and (3) does not. That is, in the situation that the Commission only sanctions after receiving the message $m = b_2$. In this case, the undisciplined government's expected utility is $EV_u = -(1-\alpha)C$ from setting its preferred deficit b_2 and $EV_u = -(b_2 - b_1)^2 - \alpha C$ if it would set the lower deficit b_1. The expected utility of setting the lower deficit is higher than the expected utility of setting the higher deficit if:

(4) $\quad b_2 < b_1 + \sqrt{C(1-2\alpha)}$

Thus, given the proposed strategies of the Commission and the disciplined government, it might pay off for the undisciplined government to set the lower deficit b_1.

The intuition of this deviation is as follows. In the initially proposed set of strategies, the undisciplined government always set its preferred higher deficit level. The likelihood that message $m = b_2$ is sent and sanctions imposed was therefore high. By setting the lower deficit, the undisciplined government deviates from its preferred budget, but also decreases the likelihood that message $m = b_2$ is sent and sanctions imposed.

Now suppose that (4) holds, and the undisciplined government sets the lower deficit b_1, like the disciplined government. How would this affect the Commission's optimal enforcement? As we outlined in (1), the Commission will impose sanctions if its expected utility from sanctioning exceeds its expected utility from not sanctioning: $EW_c > EW_{\sim c}$. However, under the newly proposed undisciplined government strategy, the Commission cannot use the message to update its beliefs over the government type in office anymore. Both the disciplined and the undisciplined government set the lower deficit b_1 and under these strategies, the message $m = b_1$ is sent with probability $1-\alpha$ and the message $m = b_2$ with probability α, regardless of the government type.

To assess this formally, consider Bayes' rule as we used before. When the Commission receives the message $m = b_1$ it updates its belief that the government in office is disciplined into: $P(j = d|m = b_1) = \frac{(1-\alpha)q}{(1-\alpha)q+(1-\alpha)(1-q)}$, which after rewriting equals the Commission's prior q. The same is true for the Commission's belief that an undisciplined government is in office after receiving $m = b_1$. Using Bayes' rule it finds that: $P(j = u|m = b_1) = \frac{(1-\alpha)(1-q)}{(1-\alpha)q+(1-\alpha)(1-q)}$, which after rewriting equals the Commission's prior $1-q$.

The same holds for the Commission beliefs after receiving the message $m = b_2$. The updated belief that a disciplined government is in office becomes $P(j = d|m = b_2) = \frac{\alpha q}{\alpha q + \alpha(1-q)}$, which equals the Commission's prior q. Finally, the updated belief that an undisciplined government is in office becomes $P(j = u|m = b_2) = \frac{\alpha(1-q)}{\alpha q + \alpha(1-q)}$, which is also equal to the Commission's prior $1 - q$. Thus, if both government types set the lower deficit b_1, the Commission cannot use the message.

Consequently, the Commission must rely on its priors to infer whether its expected utility from sanctioning exceeds its expected utility from not sanctioning. Besides, it knows that it incurs enforcement costs S when it imposes sanctions on a disciplined type and reputation costs R when it leaves the undisciplined type unsanctioned. Accordingly, it will impose sanctions if:

$$EW_c > EW_{\sim c} \quad \Rightarrow \quad P(j = u)R > P(j = d)S \quad \Rightarrow \quad (1 - q)R > qS$$

which after rewriting gives sanction condition:

$$(5) \qquad S < R\frac{1 - q}{q}$$

Note that the Commission only uses this condition if (4) holds. That is, when it knows it pays off for the undisciplined government to mimic the disciplined government and set the lower deficit b_1. The undisciplined government only considers (4) when it knows that avoiding the message $m = b_2$ pays off. That is, when (2) holds and (3) does not: When the Commission listens to the message. Moreover, if the undisciplined government sees that (5) holds and the Commission will thus sanction when (4) holds, the undisciplined government might not be interested anymore to deviate from its preferred deficit b_2, knowing that it will be sanctioned anyway. Therefore, we propose the equilibrium strategies as outlined in table 3. The corresponding (expected) utilities are outlined in table 4.

Table 3 Proposed equilibrium strategies

$b_d^* = b_1$
$b_u^* = \begin{cases} b_1 & \text{if (2) holds, (3) fails, (4) holds, and (5) fails} \\ b_2 & \text{otherwise} \end{cases}$

$$c^* = \begin{cases} \text{always} & \text{if (4) and (5) hold or (3) holds and (5) fails} \\ \text{never} & \text{if (4) holds and (5) fails or (2), (3), and (4) fail} \\ \text{message} & \text{if (2) holds and (3) and (4) fail} \end{cases}$$

Table 4 Expected utilities from the proposed equilibrium strategies

$$EV_d^* = \begin{cases} -C & \text{if (4) and (5) hold or (4) fails and (2) holds} \\ 0 & \text{if (4) holds and (5) fails or (2), (3), and (4) fail} \\ -\alpha C & \text{if (2) holds and (3) and (4) fail} \end{cases}$$

$$EV_u^* = \begin{cases} -C & \text{if (4) and (5) hold or (3) holds and (4) fails} \\ -(b_2 - b_1)^2 & \text{if (4) holds and (5) fails} \\ -(1-\alpha)C & \text{if (2) holds and (3) and (4) fail} \\ 0 & \text{if (2), (3), and (4) fail} \end{cases}$$

$$EW^* = \begin{cases} qS - (1-q)(b_1 - b_2)^2 & \text{if (4) and (5) hold or (3) holds and (4) fails} \\ -(1-q)R & \text{if (4) holds and (5) fails} \\ q\alpha S - (1-q)(b_1 - b_2)^2 - (1-q)\alpha R & \text{if (2) holds and (3) and (4) fail} \\ -(1-q)(b_1 - b_2)^2 - (1-q)R & \text{if (2), (3), and (4) fail} \end{cases}$$

In Appendix B we show that none of the actors has an incentive to deviate under the proposed strategies. We therefore conclude that the proposed strategies form the perfect Bayesian equilibrium of the game. Based on the government behavior in our equilibrium, we can distinguish two outcomes. First, there is a pooling equilibrium when both government types set deficit b_1. Second, there is a separating

equilibrium when the disciplined type sets the lower deficit b_1 and the undisciplined type the higher deficit b_2. This can be summarized as follows:

Lemma 1 *(The pooling equilibrium)*

There is a pooling equilibrium when both of the following two conditions hold. First, the Commission does not sanction while it knows that the undisciplined government sets deficit b_1, which holds if $S > R\frac{1-q}{q}$. Second, it pays off for the undisciplined government to set deficit b_1 rather than b_2 which holds if $b_2 < b_1 + \sqrt{C(1-2\alpha)}$, given that the Commission will not impose sanctions. Thus, in the pooling equilibrium both the disciplined and undisciplined governments set deficit b_1 and the Commission does not impose sanctions.

Lemma 2 *(The separating equilibrium)*

There is a separating equilibrium when one or both of the pooling conditions fails. That is, when the Commission imposes sanctions, given that it knows that the undisciplined government sets deficit b_1. This is the case when condition (5) holds: $S < R\frac{1-q}{q}$. The other pooling conditions fails when it does not pay off for the undisciplined government to set deficit b_1 instead of b_2, while it knows that the Commission would not sanction under b_1. This is the case if condition (4) fails: $b_2 > b_1 + \sqrt{C(1-2\alpha)}$. In the separating equilibrium the disciplined government sets deficit b_1, the undisciplined government sets b_2, and the Commission sanctions if both (2) and (3) hold, or if (2) holds, (3) does not, and the received message is $m = b_2$, or if (4) and (5) hold.

4.3 Theoretical findings

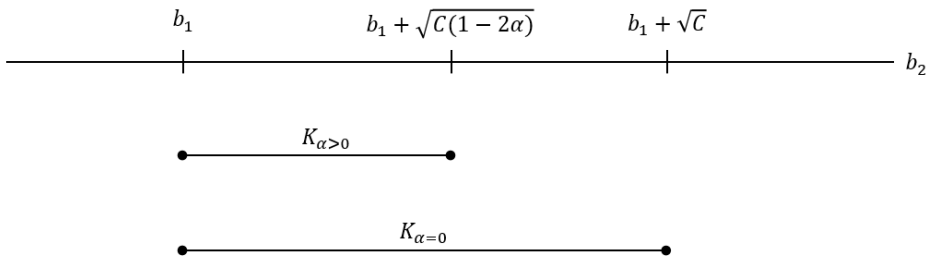
Now we have derived the equilibria of the game, we address the impact of the message's error. In particular, we assess how no or a larger error would change equilibria strategies and utilities. We find effects on government compliance and on the Commission's enforcement.

The compliance effect. A larger error reduces the range of undisciplined governments that would be willing to lower their deficit levels to avoid sanctioning. In other words, the error limits the disciplining force of the sanctions. Recall condition (4). This condition determines whether setting the lower (pooling) deficit pays off for the undisciplined government. When we would fix b_1, sanction C, and the message error α, this condition states that all undisciplined governments with a preferred deficit b_2 lower than $b_1 + \sqrt{C(1-2\alpha)}$ would prefer to pool. The threat of sanctions works effectively on these governments. On the contrary, all undisciplined governments with a preferred deficit b_2 higher than $b_1 + \sqrt{C(1-2\alpha)}$ would set the higher deficit b_2 resulting in the separating equilibrium. Yet, when the

message error becomes smaller, more undisciplined governments would fall under the first category on which the threat of sanctions has a disciplining effect. Hence, $b_1 + \sqrt{C(1-2\alpha)}$ is decreasing in α.

This is shown in figure 7. Without the message error, all the undisciplined governments in the set $K_{\alpha=0}$ would decrease their deficit due to the threat of sanctions. Due to the message error, however, only the undisciplined governments in the set $K_{\alpha>0}$ will decrease their deficit.

Figure 7 The error's effect on government compliance



The interpretation is as follows. When there is no error, the message will always been sent when governments set deficit b_2 revealing their type with certainty. This incentivizes them to set the lower deficit b_1 in order to avoid sanctions. A larger error, however, decreases the chance that the message $m = b_2$ is sent after setting deficit b_2. The error thus enables undisciplined governments to set the high deficit, without revealing their type. Therefore, the error clearly undermines the disciplining effect of fiscal rules. We summarize this in the following proposition.

Proposition 1 *(The compliance effect)*

The message's error undermines the threatening force of sanctions. When the government's high deficit is not earmarked as such, the government is not incentivized to set the low deficit to avoid sanctions. Hence, the threat of sanctions is effective for undisciplined governments with a preferred deficit lower than $b_1 + \sqrt{C(1-2\alpha)}$. *For undisciplined governments with a preferred deficit higher than* $b_1 + \sqrt{C(1-2\alpha)}$ *it is beneficial to set the preferred higher deficit, despite the threat of sanctions. Note that* $b_1 + \sqrt{C(1-2\alpha)}$ *is decreasing in* α *which means that a larger message error decreases the range of undisciplined governments on which the threat of sanctions has a disciplining effect.*

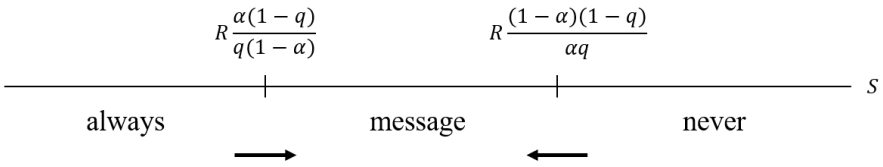
The enforcement effects. The error also affects the Commission's enforcement of the fiscal rules. In fact, the error makes the message less informative which enhances enforcement difficulties. Consider sanction conditions (2) and (3) determining whether the Commission imposes sanctions under the

separating equilibrium. The message is informative when condition (2) holds and (3) does not. In that case, the Commission will impose sanctions when the received message is $m = b_2$ and will not impose sanctions when $m = b_1$. Accordingly, by using the right-hand sides of (2) and (3) we find a range of enforcement costs for which (2) holds and (3) does not:

$$R\frac{\alpha(1-q)}{q(1-\alpha)} \leq S < R\frac{(1-\alpha)(1-q)}{\alpha q}$$

Now consider the error's effect on this range. If α increases, the left-hand side (i.e. the lower bound) will increase, whereas the right-hand (i.e. the upper bound) side will decrease. This means that the range, and thus the likelihood that the message is pivotal in the Commission's sanction decision, becomes smaller. Figure 8 pictures this graphically. More specific, an increasing left-hand side means that the likelihood that the Commission imposes sanctions when the message is $m = b_1$ increases. A decreasing right-hand side means that the likelihood that the Commission imposes sanctions when the message $m = b_2$ is sent decreases. In other words, a larger error makes the Commission ignorant regarding the message.

Figure 8 The error's effect on the equilibrium sanction strategy



This has some clear effects on the actors' utilities. First, it is beneficial for the undisciplined government but unfavorable for the disciplined government. Since the undisciplined government has in the separating equilibrium the highest probability triggering the message $m = b_2$, it benefits when the Commission ignores this message and does not sanction after receiving it. Similarly, since the disciplined government has in this equilibrium the highest probability that the message $m = b_1$ is sent, it suffers when the Commission imposes sanctions even if $m = b_1$. This is stated in the following proposition.

Proposition 2 *(The enforcement effect 1/2)*

The message's error makes the Commission ignorant regarding the message. With the error, the Commission is more inclined to impose sanctions even though the message indicates the lower deficit, while less inclined to impose sanctions while the message indicates the higher deficit. Hence, the

message is pivotal in the Commission's sanction decision when its enforcement costs S satisfy $R\frac{\alpha(1-q)}{q(1-\alpha)} \leq S < R\frac{(1-\alpha)(1-q)}{\alpha q}$, whereas a larger message error decreases this range. This favors noncompliant and harms compliant governments.

Above we found that a larger error makes it less likely that the message is pivotal for the Commission's sanction decision. However, we also find implications of the error, given that we are in the situation that the message is pivotal. Then, the error leads to enforcement mistakes: the disciplined government might be sanctioned, whereas the undisciplined government may be left unsanctioned. As a result, the error has implications for the actor's expected utilities.

First, we find that the disciplined government's expected utility in this situation is $-\alpha C$, which is decreasing in α. Second, we find that the undisciplined government's expected utility in this situation is $-(1-\alpha)C$, which is increasing in α. Third, we find that the Commission's expected utility is $-q\alpha S - (1-q)(b_1-b_2)^2 - \alpha(1-q)R$, which is also decreasing in α, hence $\frac{\partial EW}{\partial \alpha} = -qS - (1-q)R < 0$. In short, the undisciplined government benefits from the error, whereas the disciplined government and the Commission will suffer. We summarize this finding in the following proposition.

Proposition 3 *(The enforcement effect 2/2)*

In the situation that the Commission relies on the message in its enforcement, the message's error leads to inaccurate enforcement. Compliant governments might be sanctioned, whereas noncompliant governments might be left unsanctioned. Accordingly, compliant governments suffer from the error, whereas noncompliant government will benefit. The Commission will bear the corresponding reputation and enforcement costs.

In short, we find that the fiscal indicator's error decreases the disciplining force of sanctions. Undisciplined governments are less inclined to decrease their deficit levels in order to avoid sanctions, since their high deficit may not be earmarked as high. Moreover, the error makes the enforcer ignorant regarding the fiscal indicator, since the error decreases its reliability.

4.4 Assumptions

We must make some remarks on our assumed role of governments. First, in the model we assume that the uncertainty caused by the fiscal indicator is exogenous. The interpretation is that the uncertainty is the direct result of assumptions and estimates needed to derive the structural balance. In reality, however, governments are involved in the process of estimating the structural balances by providing the data to the Commission on which the estimates are based. This gives governments an opportunity to affect the

estimates and to create a biased view on their economic situation. Jonung and Larch (2006) show that France, Germany, and Italy reported structurally overoptimistic GDP growth forecasts between 1987 and 2003. Moreover, they find that these errors significantly increased the value of the CAPB, suggesting the opportunistic intentions of the governments. As main explanation they suggest that the governments try to delay economic reforms by mimicking higher CAPBs.

Second, in our model the Commission's enforcement and reputation costs are exogenous. In reality, governments might aim to affect these costs in order to influence the Commission's sanction decisions. Undisciplined governments try to increase the enforcement costs by emphasizing the economy's need for public expenditures (Wijsman and Crombez, 2017). This justifies their high deficits to a wider public and makes it politically costly for the Commission to impose sanctions. At the same time, in a multiple country setting like the EU, disciplined governments try to increase the Commission's reputation costs to make sure the Commission acts on undisciplined governments. They do so by publically emphasizing the potential negative consequences of other countries' high deficits, which makes it politically costly for the Commission to not impose sanctions.

Third, in our model the government only decides on its deficit at the beginning of the period. In reality, governments might adapt their fiscal policy during the year based on updated information on the economy. Beetsma and Giuliodori (2010) find that EU governments increase their intended deficits after new information revealed that the economy is doing better than expected. Our model does not capture this, but the reasoning is in line with our narrative. When the deficit appears (too) high after the fiscal year, governments will blame the economy.

5. The empirical model

In this section we test our theoretical findings empirically. We assess how cyclical adjustment affects the effectiveness of fiscal rules. We expand our scope to national fiscal rules in the EU, as considering the structural balance rule in the context of the SGP gives some methodological problems. We would not have a counterfactual since all European countries are subject to the SGP and we could only assess the effect of the 2005 reform in which the structural balance rule was introduced. Moreover, we assess the impact of cyclically adjusted budget rules directly on governments' fiscal performance rather than on compliance and enforcement. Due to exemption mechanisms and data limitations it is difficult to measure government compliance[16] and enforcement of fiscal rules. In line with the theoretical findings of our game-theoretical model, we have the following hypothesis:

[16] Especially when it concerns fiscal rules for regional or local governments.

Hypothesis *Fiscal rules that account for business cycle effects are less effective improving governments' fiscal performance than other fiscal rules.*

As mentioned in section 3, Debrun et al. (2008) and Ayuso-i-Casals et al. (2009) also study the effectiveness of cyclically adjusted budget rules. Our study contributes in two manners. First, we use a more advanced methodology to assess the effect of cyclically adjusted budget rules. The other papers split EU member states into two groups: one group using cyclically adjusted budget rules and the other with normal rules. They run identical regressions for both groups and consider how the effect of fiscal rules differs per group. We, on the contrary, include an index of cyclicality as independent variable which allows us to assess variance both across countries and over time. Second, our datatset covers EU-28 between 1997 and 2016 and has 525 observations, whereas the other studies have 243 and 238 observations respectively.

5.1 The basic model

We construct a panel dataset of EU-28 from 1997 till 2016 and we use a fiscal reaction function to test the effect of fiscal rules on fiscal performance. Since the work of Bohn (1998) fiscal reaction functions are used to study the fiscal sustainability of governments by assessing how governments adapt budget balances as reaction to changes in their debt levels[17]. Debrun et al. (2008) and Ayuso-i-Casals et al. (2009) use fiscal reaction functions to measure the effect of fiscal rules. Our specification becomes:

$$\mathrm{CAPB}_{i,t} = \alpha + \beta_1 \mathrm{CAPB}_{i,t-1} + \beta_2 \mathrm{FRSI}_{i,t-1} + \beta_4 \mathrm{FRSI}_{i,t-1} \times \mathrm{FRCI}_{i,t-1} + \beta_5 X_{i,t-1} + \beta_6 S_{i,t} + \beta_7 P_{i,t} + \eta_i + \mu_t + \varepsilon_{i,t}$$

We explain the variables below. Note that X, S, and P are groups of control variables.

We use the cyclically adjusted primary balance (CAPB) as percentage of potential GDP as dependent variable. Although we have shown the shortcomings of cyclical adjustment, the CAPB is the common measure for discretionary fiscal behavior as it filters out cyclical influences and the interest expenditures and is therefore fully controlled by the government (Galí and Perotti, 2003; Annett, 2006). Eyraud and Wu (2015) show that the structural balance is 25 percent more accurate measuring fiscal performance than the nominal balance. To provide robustness, we also run a set of models using the nominal budget balance in percentage of GDP as dependent variable.

We use two indices to measure the extent in which countries have cyclically adjusted budget rules. First, we use the Fiscal Rules Cyclicality Index (FRCI) as introduced by Debrun et al. (2008) to measure the extent in which individual fiscal rules take business cycle effects into account. Each fiscal rule is scored

[17] See Berti et al. (2016) for a literature review.

according the scoring scheme as constructed by Debrun et al. (2008). In short, rules that use fiscal indicators in cyclically adjusted terms or that assess compliance over multiple years receive higher scores. An annual country aggregate is calculated based on the scores of all the rules in place. The scoring is done using the European Commission's Fiscal Rules Database which contains information on 206 national fiscal rules in EU-28 between 1997 and 2016. In Appendix C we elaborate more on the construction of this index including the scoring mechanism.

Second, we use the European Commission's Fiscal Rules Strength Index (FRSI) to control for the impact of the rules. This index specifies the strength of national fiscal rules per country-year taking into account the fiscal rule coverage[18], legal base, room for revising objectives, enforcement mechanism, and resilience to shocks outside the government's control. The FRSI covers balance, expenditure, revenue, and debt rules on all government levels (total, central, regional, local) including rules imposed on social security. A higher value for the FRSI indicates stronger fiscal rules. For a more detailed explanation on the construction of the FRSI we refer to the European Commission's Fiscal Rules Database.[19]

We use both indices since the FRCI does not measure the binding character and likely impact fiscal rules have on public finances. We therefore interact the two indices to test our hypothesis. We expect that the FRSI has a positive coefficient as studies found that fiscal rules in general have an upward effect on budget balances (e.g. Bohn and Inman, 1996; Feld and Kirchgässner, 2008). By contrast, we expect the interaction coefficient to be negative as we expect cyclical adjustment to undermine the effectiveness of fiscal rules.

Figure 9 shows the unweighted cross-country average of the FRCI and FRSI between 1990 and 2015. While the FRCI was relatively stable during the 1990s, a decrease started after 2000. Over the years, fiscal rules became less cycle-friendly but after 2010 governments rely more and more on cyclically adjusted budget rules. A possible explanation for this turning point is the fiscal impact of the sovereign debt crisis. Economists and policy makers realized that deficits and debts have not been reduced sufficiently in the pre-crisis period aggravating the recession's effect on fiscal sustainability. The FRSI line shows that fiscal rules became stronger embedded over the years. Part of this trend can be explained by an increase in the number of fiscal rules. The FRSI also reflects the adoption of European fiscal rules into national law following the Six-Pack (2011) and Fiscal Compact (2012).

[18] Fiscal rule coverage refers to the percentage of government spending covered. Fiscal rules may apply to different government layers (i.e. total, central, regional, local, or social security).
[19] See European Commission (2017b)

5.2 Control variables

Output gap. We need to include a control variable to account for business cycle influences. During low economic growth governments may use deficit spending to stimulate growth. We therefore include an output gap measure as percentage of potential GDP. Using countries' own output gap would however lead to endogeneity, as our dependent variable is also expressed in potential GDP. We therefore follow Ayuso-i-Casals et al. (2009) and instrument the output gap with the weighted average output gap of the country's three largest export countries.

Debt. We control for the debt-to-GDP ratio. When public debt is higher, governments are expected to increase their fiscal efforts to decrease or stabilize their debt. Berti et al. (2016) provide an overview of literature studying governments' fiscal reaction to changing debt levels. In all sixteen studies governments were found to increase their (primary) balance when debt levels increased. We therefore expect that countries with higher debt levels have higher CAPBs. We follow the literature and include all the fiscal and economic variables with a one-year lag, including the FRCI and FRSI, as budgets are drafted prior the fiscal year (Kappeler and Välilä, 2008; European Commission, 2017a).

Euro crisis. We include a dummy variable to control for the crisis period between 2009 and 2012. This period covering both the financial and sovereign debt crisis is associated with low economic growth and banking bailouts putting downward pressure on government balances (European Commission, 2016; 2017a). We expect accordingly a negative regression coefficient.

EDP. We also include a dummy to account for the years in which EU member states are subject to an Excessive Deficit Procedure (EDP). In the context of the SGP, the Commission can propose an EDP when member states do not comply with the fiscal rules. After Council approval, EDPs give member states a deadline to end the situation of an excessive deficit. EDPs are expected to result in higher budget balances in member states.

Election. We control for parliamentary election years, since studies have shown that governments adapt fiscal policies anticipating elections. During election years, governments tend to decrease tax rates and increase expenditures to enhance their re-election chances (Clark and Hallerberg, 2000; Shi and Svensson, 2006; Alt and Lassen, 2006). We therefore expect a negative regression coefficient.

District magnitude. We control for the electoral district magnitude. We include it as the average number of parliamentary seats per electoral district. This magnitude is found to affect governments' fiscal performance in a number of ways. When the district magnitude is high there will be more political parties running for seats resulting in larger coalitions which decreases fiscal discipline (Fabrizio and

Mody, 2006). However, a higher district magnitude also loosens the connection between national politicians and their districts. This makes it easier for national politicians to demand budget discipline from sub-national governments (Foremny, 2014). Accordingly, we do not expect a certain direction of the regression coefficient sign.

Ideology. We include an index for the government's ideology. To define the government's ideology we limit ourselves to the left-right wing dimension. The ParlGov[20] database assigns index scores to political parties indicating their left-right preference. We computed the ideology score for governments based on the weighted average of the coalition partners according their respective parliamentary seats. Governments with ideology scores closer to ten are more right-wing. There is however mixed evidence on the impact of ideology on fiscal performance (Fabrizio and Mody, 2006; Debrun et al., 2008). For instance, left-wing governments are expected to have higher public expenditures but also to tax more (Volkerink and De Haan, 2001). We do not expect a certain direction of the coefficient accordingly.

Fragmentation. We control for government fragmentation. We include a variable that captures the extent in which the power within the government is centralized at one coalition party. Like Fabrizio and Mody (2006), Debrun et al. (2008), and Foremny (2014) we use the Herfindahl index which sums the squared seat shares of all coalition partners. Accordingly, single party governments have a Herfindahl index of one, whereas more fragmented governments have scores closer to zero. Roubini and Sachs (1989) found that more centralized governments have higher budget balances. When power is more dispersed within the coalition government, more constituencies can claim fiscal favors which hampers fiscal discipline. We therefore expect a positive regression coefficient.

Ideological range. In line with the fragmentation theory, we also include governments' ideological range. Using the ParlGov ideology scores, it is measured as the difference in ideology between the two most extreme coalition partners (zero if there is a one-party government). When coalition partners' vary in ideology, it is harder for governments to make budget disciplining decisions (Alesina and Drazen, 1991; Alt and Lowry, 1994). Therefore, we expect a negative regression coefficient.

5.3 Estimation methodology

We run dynamic models to account for the inertia of public finances. Budgets are generally not build from scratch each year, but are based on previous values. We therefore include the lagged dependent variable as it will have a substantial impact. We first run a set of random effects models with different set of control variables. A limitation of these models is that they do not account for the endogeneity between the dependent and its lag, but they allow us to include many control variables.

[20] Döring and Manow (2018)

Second, to account for the endogeneity between de dependent and its lag, we use the general method of moments (GMM) estimator, the mainstream estimator for dynamic panels. The GMM estimator instruments endogenous variables with their own lags. We employ both one-step difference and system GMM. The latter is less biased when the number of cross-sectional units is low (Blundell and Bond, 1998). We include different number of lags to see how the estimation results react. We report the Hansen p-value to monitor whether we include too many instruments overspecifying the variables. On the other hand, including too few instruments will fail to correct the endogenous elements. As a rule of thumb, the number of instruments should not outnumber the number of cross-sectional units (Roodman, 2009). As we have 28 individuals and 33 variables (including time dummies) overspecification is likely when we include all variables. We therefore omit the political control variables from the GMM models.

Table 5 provides an overview of our variables, data sources, and summary statistics. Table 6 shows the correlation coefficients between our variables. Except that our two alternative dependent variables are unsurprisingly high correlated, there are no other notable values.

6. Empirical results

6.1 Random effects results

Table 7 shows the estimation results of the random effects models for both dependent variables. Note that these models do not account for the dynamic panel bias. The first and fourth columns do not incorporate the cyclicality of rules, but assess the effectiveness of fiscal rules in general. The other columns include the Fiscal Rules Cyclicality Index and the interaction effect with the Fiscal Rules Strength Index. The third and sixth columns include political variables, while the second and the fifth do not. These models explain more than the half of the variance in the CAPB and nominal budget balance.

First, in line with our expectations, we find that fiscal rules are effective means to enhance fiscal discipline. The Fiscal Rules Strength Index has a positive and significant effect on the CAPB and nominal budget balance. The coefficients under the CAPB models are 0.29 and 0.30 and are significant at the 5% or 1% level. Under the second dependent variable, the coefficients range between 0.24 and 0.26 and are significant at the 10% or 5% level. These results mean that CAPBs are 0.30% of GDP higher when the Fiscal Rules Strength Index increases with one. Using its distribution, this means that moving from the 25^{th} to the 75^{th} percentile of the Fiscal Rules Strength Index increases the CAPB with 0.48% of GDP.

However, the random effects models do not provide evidence that cyclically adjusted budget rules are less effective. Under both dependent variables, the interaction effect between the Fiscal Rules Strength Index and Fiscal Rules Cyclicality Index is not significant at a meaningful level. Also when the political variables are added in the third and sixth columns, we do not find significant results. Moreover, under the CAPB models the coefficients are close to zero. Also the direct effect of the Fiscal Rules Cyclicality Index on our dependent variables is not significant at a meaningful level.

As regards the control variables, the findings are mixed. The coefficients of the output gap are not statistically significant at a meaningful level and differ between the models. Under the first dependent variable, the coefficients are negative ranging between -0.06 and -0.08. Under the second dependent variable the coefficients are 0.03 or 0.04. It is difficult to draw conclusions from these findings, as the coefficients do not show clear pro or countercyclical fiscal policies. For instance, it is possible that pro and countercyclical forces annul each other. The random effect models suggest that member states with higher debt level have lower budget balances. The debt coefficients under the CAPB are -0.01 and statistically significant at the 5% or 10% level. Also under the budget balance models, the debt coefficients are negative, but close to zero and not statistically significant. These findings contrast our expectation that member states with more debt would strengthen fiscal effort increasing budget balances.

Also the results of the dummy variables are mixed. Against our expectation, the euro crisis dummies are positive in five of the models, but only significant at the 5% level in two of these models. The euro crisis coefficients differ substantially in the models containing political variables. The results related to the election dummy suggests government budgets are lower during election years. The election coefficients are negative in all models and statistically significant at the 5% level in two models. Under the CAPB models, the coefficients range between -0.42 and -0.64, and under the budget balance models between -0.76 and -1.01. Using the significant results in column four and five, we find that nominal budget balances are 1% of GDP lower during election years. The EDP dummies do not provide clear results. In five models the coefficients are negative, but close to zero. In none of the models the EDP dummy is significant at a meaningful level.

The results regarding the political variables are broadly as expected. The district magnitude decreases CAPBs and nominal budget balances. The coefficients are significant at the 1% level, but are also close to zero. As expected, also fragmentation hampers fiscal discipline. The coefficients are -1.97 and -1.91 and are statistically significant at the 5% level. On the contrary, there seems no effect of government ideology and the ideological range, as no significant results are found.

6.2 GMM results

Table 8 shows the regression outcomes of the difference and system GMM models for both dependent variables. Per dependent variable there are three difference GMM and two system GMM models. Columns one and six do not test the effect of cyclically adjusted budget rules but only the effect of fiscal rules in general. The other columns vary regarding the number of lags that we allow to instrument the endogenous variables. However, using different lags does not change the results.

We again find that fiscal rules increase fiscal discipline. The Fiscal Rules Strength Index coefficients are positive in all GMM models, but only under system GMM significant at the 5% or 1% level. The coefficients under the first dependent variable range between 0.32 and 0.34, and under the second dependent variable between 0.25 and 0.26. Despite having both difference and system GMM, and including different numbers of lags, these coefficients are stable and similar to the coefficients under the random effects models. The Fiscal Rules Strength Index coefficient of 0.34 in the system GMM models under the first dependent means that the CAPB increases with 0.54% of GDP moving from the 25^{th} to 75^{th} percentile of the Fiscal Rules Strength Index.

Also under the GMM models no evidence is found that cyclically adjusted budget rules are less effective. The coefficients of the interaction between the Fiscal Rules Strength Index and the Fiscal Rules Cyclicality Index vary between the models and are not significant. Under the CAPB models, the coefficients are positive (0.02 and 0.07), but the system GMM models under the nominal budget balance models report negative coefficients (-0.02). Also the direct effect from the Fiscal Rules Strength Index on the budget indicators is not statistically significant. As both the random effect models and the GMM models do not find any effect related to the Fiscal Rules Cyclicality Index, we do not find evidence for our hypothesis that fiscal rules that account for business cycle effects are less effective.

As under the random effects models, the coefficients of the output gap do not provide a clear effect. Under the CAPB models they are slightly negatives, whereas under the nominal balance models they are slightly positive. In none of the models the coefficients are significant at a meaningful level. These results do not provide an indication for pro- or countercyclical fiscal policies. Also the results related to debt levels are ambiguous. The difference GMM models find that the CAPB and nominal balances are increasing in debt levels. The coefficients are 0.09 and 0.10 and are statistically significant at the 1% level. This indicates, as expected, that governments adopt fiscal consolidation measures when debt levels increase. However, the system GMM models point towards a negative effect between debt levels and budgets.

Regarding the dummy variables, the results are mixed. The euro crisis dummies are dropped due to collinearity in most of the GMM models. In only two difference GMM models under the second dependent variable the euro crisis dummies are sustained. In these models the coefficients are positive, in contrast to our expectation, but not significant at a meaningful level. As under the random effects models, governments seem to have lower CAPBs and nominal balances in election years. Under the CAPB models, the election coefficients range between -0.17 and -0.60, but are not statistically significant. Under the second dependent variable, the coefficients range between -0.31 and -0.90. Only the system GMM models are significant, but just at the 10% level. Finally, the EDP dummies under the GMM models show that member states improve their public finances under EDPs. Under the first dependent variable, all EDP coefficients are positive, ranging between 0.16 and 0.46, but none of them is significant. Under the second dependent variable, the coefficients range between 0.14 and 0.57 and are significant at the 5% under the difference GMM. These result show that nominal balances are about 0.57% of GDP higher during EDP years.

Although we included the lagged dependent to account for autocorrelation, the AR(2) test does not strongly reject the null-hypothesis of no second-order autocorrelation. Depending on the model, the p-value ranges between 0.075 and 0.143. We therefore run another set of models including the dependent variable with two lags while leaving the other variables unchanged. Table 9 depicts the results of these models. The results are broadly similar compared to our initial GMM models. Fiscal rules seem to increase fiscal discipline, but cyclical adjustment does not influence this effect. Elections worsen public finances, while balances are higher during EDPs. In line with our expectations, CAPBs and nominal budget balances were significantly lower during the euro crisis. While the results concerning debt-to-GDP levels were ambiguous under the random effects and initial GMM models, table 8 shows the positive effect of debt on budget balances.

In short, our estimation models do not provide evidence that cyclically adjusted budget rules are less effective than simpler fiscal rules. This contrasts the finding of Debrun et al. (2008). They separated EU member states into two groups: one group of member states scoring high on the Fiscal Rules Cyclicality Index and the other group scoring low. They found that the Fiscal Rules Strength Index had a stronger effect in the low scoring countries. We, on the contrary, modelled an interaction effect between the two indices. This maintained the continuous nature of the Fiscal Rules Cyclicality Index, instead of using country groups. However, under this methodology, the effect found by Debrun et al. does not sustain. A possible explanation that we do not find evidence for our hypothesis is the use of multiple indices. The indices are based on scoring mechanisms following checklists. These are carefully constructed, but are unavoidable subject to noise.

7. Conclusion

This paper builds on the fiscal policy literature assessing the effectiveness of cyclically adjusted budget rules. Focusing on the EU's structural balance rule, the literature has shown that business cycle estimates lead to frequent revisions of the fiscal indicator. We argued that this creates uncertainty since governments do not fully control the fiscal indicator used for the structural balance rule. Although the literature stressed this problem, it addressed the implications for fiscal rules' effectiveness rather limited. We therefore assessed theoretically how the uncertainty affects the incentives politicians have to comply with fiscal rules and the incentives of the enforcer. We tested our theoretical findings empirically.

We found that the uncertainty stemming from cyclical adjustment undermines the disciplining force of fiscal rules. In particular, we found compliance and enforcement effects. First, when governments' noncompliant fiscal policies are not labelled as noncompliant, they are not incentivized to improve their fiscal policies. Second, the enforcer ignores the fiscal indicator when this indicator is unreliable. As a result, the enforcer will impose sanctions while fiscal policies are labelled as compliant, or will omit sanctions while fiscal policies are labelled as noncompliant. It might thus happen that compliant governments are sanctioned, while noncompliant governments are left unsanctioned. Using a panel dataset of EU-28 between 1997 and 2016, we tested the theoretical findings empirically, but we did not find evidence that cyclically adjusted budget rules are less effective.

7.1 The structural balance and SGP reform

The structural balance rule is adopted in 2005 to make the SGP 'smart', meaning 'binding' regardless the economic times. However, smartness came at the expense of complexity (i.e. the broader public has difficulties understanding the rules) and uncertainty (i.e. the structural balance is not fully under control of the government). The experience since the introduction of the structural balance rule led economists and policy makers to outline suggestions for reform. Eyraud and Wu (2015) list five possible ways forward: (*i*) Methodological improvements, (*ii*) explicitly account for the bias ex-ante, (*iii*) using a control account recording ex-post deviations, (*iv*) replacing the structural balance with an indicator mimicking its properties, or (*v*) abandoning the structural balance rule.

The first two alternatives have been Europe's way to deal with the structural balance's uncertainty. For decades, economists both inside and outside academia aim to improve cyclical adjustment with methodological improvements[21] (e.g. Bouthevillain et al., 2001; Denis et al., 2006). Based on OECD research[22] the Commission introduced a new methodology to calculate the cyclically adjusted balance

[21] In our theoretical model a reduction of error α.
[22] Girouard and André (2005)

in June 2012 (Mourre et al., 2013) and also in recent years institutions like the European Commission[23], IMF[24], and OECD[25] continued research aiming for methodological improvements. Also the permanent Output Gaps Working Group of the Economic Policy Committee aims to improve the estimations of potential output.

The third alternative, using a control account, is proposed by Feld et al. (2018). The idea is that structural balance revisions, together with deviations from the structural balance rule, are recorded in a fictitious account. This mitigates the uncertainty as upward revisions will compensate downward revisions over time. If this fictitious account is used to assess fiscal discipline, monitoring becomes less dependent on a single year. However, it would not erase the uncertainty enhanced by the revisions completely. Christofzik et al. (2018) show the origin of fluctuations in this account for EU member states if it had been implemented between 2013 and 2017. Especially in Spain and Italy net revisions of the structural balance account for a substantial share in the change in the fictitious account. For Spain 2014, the revisions of past structural balances is larger than the real-time violation of the structural balance in 2014. Another caveat is that a fictitious account is still difficult to understand for the broader public.

In line with the fourth and fifth alternatives, however, most SGP reform proposals suggest abandoning the structural balance rule and using an expenditure (growth) rule instead (European Fiscal Board, 2018, Bénassy-Quéré et al., 2018; Darvas, Martin, Ragot, 2018, Christofzik et al., 2018). First, unlike the structural balance, public expenditures are observable and controlled by the government (Darvas, Martin, and Ragot, 2018). Compared to government revenues and budget balances, expenditures are less sensitive to the economic cycle (Christofzik et al.,2018). Using public expenditures to monitor fiscal discipline is accordingly more efficient. Especially when one subtracts some observable cyclical elements from the public expenditure level, cyclical forces can be filtered out. For instance, Darvas, Martin, and Ragot (2018) suggest using nominal expenditures net of interest payments and unemployment spending.

European policy makers have been discussing the problems with the structural balance rule, but no concrete solutions materialized. In November 2016, the Economic and Financial Committee, the EU's committee to promote coordination among member states, recommended a stronger focus on the existing expenditure rule under the Preventive Arm to improve the predictability and transparency of the SGP[26]. In its May 2017 Reflection Paper on the deepening of the economic and monetary union, the

[23] See for instance Hristov, Raciborski, and Vandermeulen (2017)
[24] See for instance Alichi (2015)
[25] See for instance Turner et al. (2016)
[26] See Opinion of the Economic and Financial Committee "Improving the predictability and transparency of the SGP: A stronger focus on the expenditure benchmark in the Preventive Arm" on 29 November 2016.

Commission indicated that a simplification of the fiscal rules is on the agenda for 2020-2025, but no concrete proposals were made yet (European Commission, 2017c).

Figure 9 The unweighted cross-country average of the FRCI and FRSI (EU-28)

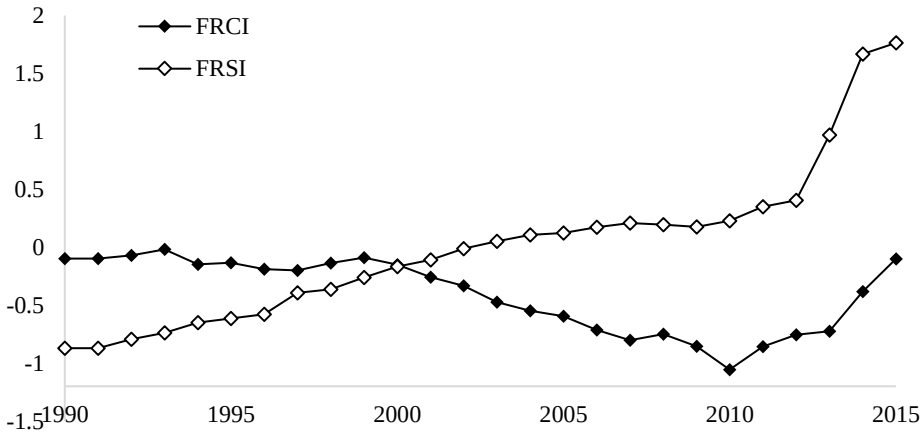


Table 5 Data sources and summary statistics

Variable	Source	Mean	Std. dev.	Min	Max
CAPB	AMECO	-5.214	4.118	-31.640	4.599
Balance	AMECO	-2.670	3.505	-32.051	6.855
Fiscal Rules Strength Index	European Commission	0.269	1.014	-0.991	3.525
Fiscal Rules Cyclicality Index	European Commission Own calculations	-0.515	1.123	-4.92	3.609
Foreign output gap	OECD	-0.613	1.933	-7.652	7.227
Debt	AMECO	55.580	32.272	3.664	180.833
Euro crisis		0.2	0.4	0	1
Election	CKS (2015)	0.866	0.341	0	1
EDP		0.361	0.481	0	1
District magnitude	CKS (2015) Own calculations	15.209	26.813	1	150
Ideology	ParlGov Own calculations	5.419	1.533	1.0526	8.6842
Fragmentation	ParlGov Own calculations	0.635	0.260	0.181	1
Ideological range	ParlGov Own calculations	2.044	1.729	0	7.6316

N = 28, T = 20 (1997 – 2016)
CKS (2015) refers to Cruz, Keefer, and Scartascini (2015)

Table 6 Correlation matrix

		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
(1)	CAPB	1.000												
(2)	Balance	0.854	1.000											
(3)	FRSI	0.362	0.229	1.000										
(4)	FRCI	0.082	0.093	-0.032	1.000									
(5)	Output gap	0.090	0.397	-0.162	0.097	1.000								
(6)	Debt	-0.525	-0.399	-0.014	0.023	-0.240	1.000							
(7)	Euro crisis	-0.197	-0.375	-0.006	-0.135	-0.428	0.134	1.000						
(8)	Election	-0.065	-0.103	-0.101	-0.100	0.116	-0.115	0.163	1.000					
(9)	EDP	-0.315	-0.426	0.044	-0.104	-0.305	0.410	0.451	0.158	1.000				
(10)	District magnitude	-0.038	-0.083	0.005	-0.138	-0.032	-0.114	0.019	0.024	0.053	1.000			
(11)	Ideology	0.119	0.051	0.018	0.040	0.006	-0.067	0.094	0.124	-0.034	-0.019	1.000		
(12)	Fragmentation	-0.265	-0.218	-0.030	0.054	-0.008	0.179	-0.031	-0.031	0.113	-0.189	-0.145	1.000	
(13)	Ideological range	0.084	0.063	-0.060	-0.031	0.007	-0.059	-0.047	0.016	-0.120	0.092	0.011	-0.682	1.000

Table 7 Estimation results: random effects (EU-28, 1997-2016)

Dep. variable:	CAPB			Balance		
	1	2	3	4	5	6
Lagged dependent	0.77***	0.77***	0.74***	0.72***	0.72***	0.70***
	(12.93)	(12.76)	(12.41)	(13.90)	(13.82)	(14.22)
Lagged FRSI	0.29**	0.30**	0.30***	0.26**	0.24*	0.25**
	(2.44)	(2.28)	(2.57)	(2.01)	(1.89)	(2.10)
Lagged FRCI		0.03	0.04		0.04	0.06
		(0.23)	(0.35)		(0.37)	(0.48)
Lagged FRSI × Lagged FRCI		0.01	0.00		-0.04	-0.05
		(0.13)	(0.00)		(-0.61)	(-0.83)
Lagged output gap	-0.06	-0.07	-0.08	0.03	0.03	0.04
	(-0.97)	(-1.03)	(-1.11)	(0.41)	(0.45)	(0.52)
Lagged debt	-0.01**	-0.01**	-0.01*	-0.00	-0.00	-0.00
	(-2.17)	(-2.18)	(-1.89)	(-0.97)	(-0.98)	(-0.46)
Euro crisis (dummy)	0.86**	0.90**	0.29	0.37	0.38	-0.26
	(1.96)	(1.96)	(0.51)	(0.78)	(0.77)	(-0.42)
Election (dummy)	-0.63	-0.64	-0.42	-1.00**	-1.01**	-0.76
	(-1.60)	(-1.61)	(-0.91)	(-2.35)	(-2.36)	(-1.54)
EDP (dummy)	-0.03	-0.04	0.05	-0.11	-0.11	-0.05
	(-0.12)	(-0.14)	(0.18)	(-0.45)	(-0.45)	(-0.17)
District magnitude			-0.00***			-0.01***
			(-2.69)			(-3.37)
Ideology			0.07			0.02
			(1.13)			(0.39)
Fragmentation			-1.97**			-1.91**
			(-2.12)			(-2.01)
Ideological range			-0.15			-0.17
			(-0.98)			(-1.02)
Constant	0.23	0.22	1.46	0.36	0.39	1.90*
	(0.73)	(0.70)	(1.32)	(0.98)	(1.06)	(1.68)
Observations	525	525	520	525	525	520
R^2 within	0.51	0.51	0.52	0.57	0.57	0.57

Note: The columns contain estimation results of the random effects models. T-statistics are in parentheses, and *, **, *** denote significance at the 10, 5, and 1% level respectively. Time dummies are not shown due to space limitations.

Table 8 Estimation results: GMM (EU-28, 1997-2016)

Dep. variable:	CAPB					Balance				
	Difference	Difference	Difference	System	System	Difference	Difference	Difference	System	System
Lagged dependent	0.45***	0.45***	0.44***	0.73***	0.71***	0.48***	0.45***	0.46***	0.70***	0.70***
	(4.66)	(4.58)	(4.40)	(11.76)	(9.27)	(5.76)	(5.28)	(5.44)	(13.14)	(13.27)
Lagged FRSI	0.32	0.33	0.33	0.34***	0.34**	0.26	0.26	0.26	0.25**	0.25**
	(1.38)	(1.50)	(1.50)	(2.56)	(2.55)	(1.04)	(1.16)	(1.16)	(1.99)	(1.98)
Lagged FRCI		-0.40	-0.39	-0.03	-0.03		-0.41	-0.42	-0.03	-0.03
		(-1.20)	(-1.18)	(-0.21)	(-0.19)		(-1.22)	(-1.23)	(-0.19)	(-0.19)
Lagged FRSI × lagged FRCI		0.07	0.07	0.02	0.02		0.03	0.03	-0.02	-0.02
		(0.42)	(0.44)	(0.20)	(0.24)		(0.20)	(0.19)	(-0.28)	(-0.28)
Lagged output gap	-0.05	-0.02	-0.03	-0.06	-0.06	0.08	0.10	0.10	0.04	0.04
	(-0.29)	(-0.13)	(-0.17)	(-0.66)	(-0.66)	(0.50)	(0.63)	(0.62)	(0.44)	(0.44)
Lagged debt	0.09***	0.09***	0.10***	-0.01***	-0.01***	0.09***	0.09***	0.09***	-0.00	-0.00
	(3.61)	(3.56)	(3.57)	(-2.82)	(-2.66)	(5.07)	(4.98)	(4.98)	(-0.72)	(-0.81)
Euro crisis (dummy)	Omitted	Omitted	Omitted	Omitted	Omitted	Omitted	0.50 (0.68)	0.51 (0.69)	Omitted	Omitted
Election (dummy)	-0.17	-0.27	-0.18	-0.54	-0.60	-0.31	-0.37	-0.32	-0.90*	-0.89*
	(-0.24)	(-0.40)	(-0.26)	(-0.97)	(-1.18)	(-0.53)	(-0.64)	(-0.55)	(-1.67)	(-1.68)
EDP (dummy)	0.45	0.42	0.46	0.19	0.16	0.57**	0.52**	0.57**	0.14	0.15
	(1.58)	(1.47)	(1.62)	(0.77)	(0.61)	(2.06)	(2.04)	(2.13)	(0.65)	(0.68)
Constant				-0.71	-1.70				-0.92	-0.92
				(-1.13)	(-1.14)				(-0.76)	(-0.77)
Observations	497	497	497	525	525	497	497	497	525	525
Lags	6	4	5	3	4	7	4	5	3	4
No. of instruments	29	29	30	30	31	30	29	30	30	31
AR(1) (p-value)	0.004	0.004	0.004	0.006	0.004	0.020	0.20	0.020	0.012	0.010
AR(2) (p-value)	0.122	0.143	0.140	0.130	0.135	0.088	0.096	0.095	0.077	0.075
Hansen (p-value)	0.721	0.730	0.833	0.818	0.967	0.571	0.754	0.997	0.435	0.769

Note: The columns contain estimation results of the GMM models. T-statistics are in parentheses, and *, **, *** denote significance at the 10, 5, and 1% level respectively. Time dummies are not shown due to space limitations. The AR(1) and AR(2) test whether residuals are autocorrelated in the first and the second lag. To avoid overspecification in the GMM model, we have collapsed the matrix of instruments and restricted the set of internal instruments. The Hansen test confirms that the instruments as groups are exogenous.

Table 9 Estimation results: GMM (EU-28, 1997-2016)

Dep. variable:	CAPB					Balance				
	Difference	Difference	Difference	System	System	Difference	Difference	Difference	System	System
Lagged dependent	0.35*** (3.22)	0.35*** (3.10)	0.34*** (3.06)	0.70*** (9.38)	0.70*** (9.40)	0.44*** (5.37)	0.42*** (4.67)	0.42*** (4.67)	0.69*** (13.55)	0.69*** (13.89)
Second lagged dependent	0.12** (2.32)	0.12** (2.23)	0.12** (2.27)	0.09 (1.37)	0.09 (1.35)	0.19*** (3.36)	0.18*** (3.15)	0.18*** (3.12)	0.14** (2.56)	0.15*** (2.57)
Lagged FRSI	0.28 (1.24)	0.25 (1.15)	0.27 (1.20)	0.26* (1.89)	0.26* (1.89)	0.09 (0.38)	0.06 (0.27)	0.07 (0.31)	0.11 (0.94)	0.10 (0.88)
Lagged FRCI		-0.37 (-1.13)	-0.36 (-1.09)	-0.03 (-0.20)	-0.03 (-0.20)		-0.41 (-1.16)	-0.39 (-1.11)	-0.04 (-0.21)	-0.04 (-0.22)
Lagged FRSI × lagged FRCI		0.08 (0.53)	0.08 (0.53)	0.01 (0.18)	0.01 (0.18)		0.06 (0.36)	0.05 (0.32)	-0.02 (-0.28)	-0.02 (-0.27)
Lagged output gap	-0.09 (-0.48)	-0.07 (-0.38)	-0.09 (-0.45)	-0.09 (-0.98)	-0.09 (-0.97)	0.00 (0.01)	0.03 (0.13)	0.02 (0.09)	-0.04 (-0.42)	-0.04 (-0.44)
Lagged debt	0.11*** (4.09)	0.11*** (4.01)	0.12*** (4.05)	-0.00 (-0.64)	-0.00 (-0.63)	0.10*** (4.54)	0.11*** (5.25)	0.11*** (5.21)	0.00 (1.00)	0.00 (1.09)
Euro crisis (dummy)	-1.46 (-0.81)	-3.57** (-2.25)	-0.37 (-0.41)	0.02 (0.05)	-1.54 (-0.99)	-4.84*** (-6.31)	-5.11*** (-6.19)	-2.61*** (-3.02)	-1.29 (-1.19)	-4.18*** (-6.10)
Election (dummy)	-0.36 (-0.40)	-0.34 (-0.39)	-0.34 (-0.40)	-0.79 (-1.01)	-0.79 (-1.01)	-0.92 (-1.27)	-0.91 (-1.36)	-0.91 (-1.35)	-1.37** (-2.15)	-1.37** (-2.16)
EDP (dummy)	0.39 (1.42)	0.44 (1.61)	0.43 (1.56)	0.23 (0.78)	0.23 (0.78)	0.53** (2.44)	0.61** (2.55)	0.59*** (2.58)	0.33 (1.63)	0.34 (1.63)
Constant				-1.49* (-1.67)	0.04 (0.11)				0.40 (0.88)	0.40 (0.89)
Observations	470	470	470	498	498	470	470	470	498	498
Lags	10	6	7	4	5	10	6	7	3	4
No. of instruments	33	31	32	32	33	33	31	32	31	32
AR(1) (p-value)	0.006	0.010	0.009	0.004	0.004	0.011	0.018	0.018	0.008	0.008
AR(2) (p-value)	0.714	0.651	0.613	0.055	0.052	0.690	0.691	0.669	0.059	0.067
AR(3) (p-value)	0.951	0.920	0.928	0.650	0.649	0.498	0.397	0.402	0.107	0.107
Hansen (p-value)	0.981	0.752	0.898	0.657	0.836	0.801	0.377	0.782	0.431	0.923

Note: The columns contain estimation results of the GMM models. T-statistics are in parentheses, and *, **, *** denote significance at the 10, 5, and 1% level respectively. Time dummies are not shown due to space limitations. The AR(1), AR(2), and AR(3) test whether residuals are autocorrelated in the first, second, and third lag. To avoid overspecification in the GMM model, we have collapsed the matrix of instruments and restricted the set of internal instruments. The Hansen test confirms that the instruments as groups are exogenous.

8. References

Alesina, A. and Ardagna, S. (2010), "Large Changes in Fiscal Policy: Taxes versus Spending", *Tax Policy and the Economy*, 24 (1), pp. 35-68

Alesina, A. and Drazen, A. (1991), "Why Are Stabilizations Delayed?", *The American Economic Review*, 81 (5), pp. 1170-1188

Alesina, A. and Perotti, R. (1995), "Fiscal expansions and adjustments in OECD countries", *Economic Policy*, 10 (21), pp. 205-248

Alichi, A. (2015), "A New Methodology for Estimating the Output Gap in the United States", *IMF Working Paper No. 15/144*

Alt, J. E. and Lassen, D. D. (2006), "Transparency, Political Polarization, and Political Budget Cycles in OECD Countries", *American Journal of Political Science*, 50 (3), pp. 530-550

Alt, J. E., and Lowry, R. C. (1994), "Divided Government, Fiscal Institutions, and Budget Deficits: Evidence from the States", *American Political Science Review*, 88 (4), pp. 811-828

Annett, A. (2006), "Enforcement and the Stability and Growth Pact: How Fiscal Policy Did and Did not Change Under Europe's Fiscal Framework", *IMF Working Paper 06/116*

Ayuso-i-Casals, J., González Hernández, D., Moulin, L., and Turrini, A. (2009), "Beyond the SGP – Features and effects of EU national-level fiscal rules", in: "*Policy Instruments for Sound Fiscal Policies. Fiscal Rules and Institutions*", Palgrave Macmillan, London, pp. 204-240

Beetsma, R. and Giuliodori, M. (2010), "Fiscal adjustment to cyclical developments in the OECD: an empirical analysis based on real-time data", *Oxford Economic Papers*, 62, pp. 419-441

Bénassy-Quéré, A., Brunnermeier, M., Enderlein, H., Farhi, E., Fratzscher, M., Fuest, C., Gourinchas, P., Martin, P., Pisani-Ferry J., Rey, H., Schnabel, I., Véron, N., Weder di Mauro, B., and Zettelmeyer,

J. (2018), “Reconciling risk sharing with market discipline: A constructive approach to euro area reform”, *Centre for Economic Policy Research, Policy Insight No. 91*

Berti, K., Colesnic, E., Desponts, C., Pamies, S., and Sail, E. (2016), "Fiscal Reaction Functions for European Union Countries", *European Economy Discussion Paper 028 April 2016*

Blundell, R. and Bond, S. (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87, pp. 115-143

Bohn, H. (1998), “The behavior of US public debt and deficits”, *Quarterly Journal of Economics*, 113, pp. 949-963

Bohn, H. and Inman, R.P. (1996), “Balanced-budget rules and public deficits: evidence from the U.S. states”, *Carnegie-Rochester Conference Series on Public Policy*, 45, pp. 13-76

Bouthevillain, C., Cour-Thimann, P., Van den Dool, G., Hernandez de Cos, P., Langenus, G., Mohr, M., Momigliano, S., Tujula M. (2001), “Cyclically Adjusted Budget Balances: An Alternative Approach”, *ECB Working Paper Series* 77

Buti, M. (2016), “What Future for Rules-Based Fiscal Policy?”, in: Rajan, R., Rogoff, K. Summer, L.H., and Blanchard, O. (2016), “*Progress and Confusion: The State of Macroeconomic Policy*”, MIT Press 2016, Cambridge, Massachusetts, pp. 183-195

Chaudoin, S. (2014), “Audience Features and the Strategic Timing of Trade Disputes”, *International Organization*, 68 (3), pp. 877-911

Christofzik, D., Feld, L. P., Reuter, W. H., and Yeter, M. (2018), “Uniting European Fiscal Rules: How to Strengthen the Fiscal Framework”, *Working Paper 04/2018*, German Council of Economic Experts

Cimadomo, J. (2012), “Fiscal Policy in Real Time”, *The Scandinavian Journal of Economics*, 114 (2), pp. 440-465

Cimadomo, J. (2014), “Real-Time Data and Fiscal Policy Analysis: A Survey of the Literature”, *Journal of Economic Surveys*, 30 (2), pp. 302-326

Claeys, G., Darvas, Z., and Leandro, A. (2016), "A proposal to revive the European fiscal framework", *Bruegel Policy Contribution,* 2016/07

Clark, W. R. and Hallerberg, M. (2000), "Mobile Capital, Domestic Institutions, and Electorally Induced Monetary and Fiscal Policy", *American Political Science Review,* 94 (2), pp. 323-346

Cordes, T., Kinda, T., Muthoora, P., and Weber, A. (2015), "Expenditure Rules: Effective Tools for Sound Fiscal Policy?", *IMF Working Paper 15/29*

Costantini, O. (2015), "The Cyclically Adjusted Budget: History and Exegesis of a Fateful Estimate", *Institute for New Economic Thinking Working Paper No. 24*

Cruz, C., Keefer, P., and Scartascini, C. (2015), "The Database of Political Institutions 2015", *World Bank Economic Review*

Darvas, Z., Martin, P., and Ragot, X. (2018), "European Fiscal Rules Require a Major Overhaul", *Les notes du conseil d'analyse économique,* French Council of Economic Analysis

Debrun, X., Moulin, L., Turrini, A., Ayuso-i-Casals, J., and Kumar, M. S. (2008), "Tied to the mast? National fiscal rules in the European Union", *Economic Policy,* April 2008, pp. 297-362

Denis, C., Grenouilleau, D., Mc Morrow, K., and Röger, W. (2006), "Calculating Potential Growth Rates and Ouput Gaps – A Revised Production Function Approach", *European Economy. Economic Papers 247*

Döring, H. and Manow, P. (2018), "*Parliaments and governments database (ParlGov): Information on parties, elections, and cabinets in modern democracies*"

European Commission (2016), "*Specifications on the implementation of the Stability and Growth Pact and Guidelines on the format and content of Stability and Convergence Programmes*"

European Commission (2017a), "*European Economy. Public finances in EMU 2017*"

European Commission (2017b), "*Fiscal Rules Database New Methodology*"

European Commission (2017c), "*Reflection paper on the deepening of the economic and monetary union*"

European Commission (2018), "*Vade Mecum on the Stability & Growth Pact. 2018 edition*"

European Fiscal Board (2018), "*Annual Report 2018*", 28 September 2018

Eyraud, L. and Wu, T. (2015), "Playing by the Rules: Reforming Fiscal Governance in Europe", *IMF Working Paper 15/67*

Fabrizio, S. and Mody, A. (2006), "Can budget institutions counteract political indiscipline?", *Economic Policy*, October 2006, pp. 689-739

Feld, L.P. and Kirchgässner, G. (2008), "On the effectiveness of debt brakes: the Swiss experience", In: Neck, R. and Sturm, J. (2008), "*Sustainability of Public Debt*", The MIT Press, Cambridge

Feld, L.P., Schmidt, C., Schnabel, I., and Wieland, V. (2018), "*Refocusing the European fiscal framework*", Vox.eu, 12 September 2018

Foremny, D. (2014), "Sub-national deficits in European countries: The impact of fiscal rules and tax autonomy", *European Journal of Political Economy,* 34, pp. 86-110

Galí, J. and Perotti, R. (2003), "Fiscal policy and monetary integration in Europe", *Economic Policy,* 18 (37), pp. 533-572

Girouard, N. and André, C. (2005), "Measuring cyclically-adjusted budget balances for OECD countries", *OECD Economics Department Working Papers No. 434*

Hristov, A., Raciborski, R., and Vandermeulen, V. (2017), "Assessment of the Plausibility of the Output Gap Estimates", *European Economy. Economic Brief 023*

Hughes Hallett, A, Kattai, R., and Lewis, J. (2012), "How Reliable are Cyclically Adjusted Budget Balances in Real Time?", *Contemporary Economic Policy,* 30 (1), pp. 75-92

Jonung, L. and Larch, M. (2006), "Improving fiscal policy in the EU: the case for independent forecasts", *Economic Policy,* July 2006, pp. 491-534

Kappeler, A. and Välilä, T. (2008), "Fiscal federalism and the composition of public investment in Europe", *European Journal of Political Economy,* 24, pp. 562-570

Larch, M. and Turrini, A. (2010), "The Cyclically Adjusted Budget Balance in EU Fiscal Policymaking", *Intereconomics: Review of European Economic Policy*, 45 (1), pp. 48-60

Morris, R., Ongena, H., and Schuknecht, L. (2006), "The Reform and Implementation of the Stability and Growth Pact", *European Central Bank Occasional Paper Series No. 47 / June 2006*

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

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

Orphanides, A. and Van Norden, S. (2002), "The Unreliability of Output-Gap Estimates in Real Time", *The Review of Economics and Statistics,* 84 (4), pp. 569-583

Roodman, D. (2009), "How to do xtabond2: An introduction to Difference and System GMM in Stata", *Stata Journal,* 9, pp. 86-136

Roubini, N. and Sachs, J. D. (1989), "Political and Economic Determinants of Budget Deficits in the Industrial Democracies", *European Economic Review,* 33, pp. 903-938

Schaechter, A., Kinda, T., Budina, N., and Weber, A. (2012), "Fiscal Rules in Response to the Crisis. Toward the "Next-Generation" Rules. A New Dataset", *IMF Working Paper 12/187*

Schuknecht, L. (2004), "EU Fiscal Rules. Issues and Lessons from Political Economy", *ECB Working Paper Series No. 421*

Shi, M. and Svensson, J. (2006), "Political budget cycles: Do they differ across countries and why?", *Journal of Public Economics,* 90 (8-9), pp. 1367-1389

Turner, D., Cavalleri, M. C., Guillemette, Y., Kopoin, A., Ollivaud, P., and Rusticelli, E. (2016), "An investigation into improving the real-time reliability of OECD output gap estimates", *OECD Economics Department Working Papers No. 1294*

Volkerink, B. and De Haan, J. (2001), "Fragmented government effects on fiscal policy: New evidence", *Public Choice*, 109, pp. 221-242

Wijsman, S. and Crombez, C. (2017), "The Political Economy of Enforcing Fiscal Rules", *KU Leuven MSI Working Paper 1701*

Wolff, G. (2017), "Beyond the Juncker and Schäuble visions of euro-area governance", *Bruegel Policy Brief Issue 6*

9. Appendix

Appendix A: Equilibrium derivation

To infer whether this is an equilibrium, we check whether the government and/or the Commission has an incentive to deviate from the proposed strategy. First, consider the case in which both (2) and (3) hold. The disciplined government has no incentive to set another budget than its preferred b_1. When it would set the higher budget b_2, it would obtain $V_d = -(b_1 - b_2)^2 - C$, which is lower than $-C$, its utility after setting b_1.

Also the undisciplined government has no incentive to set another budget than its preferred b_2. When it would set the lower budget b_1 instead, it obtains $V_u = -(b_2 - b_1)^2 - C$, which is lower than $-C$, which is obtained after setting b_2.

Moreover, the Commission has no incentive to deviate from its proposed strategy. Suppose it would never sanction while both (2) and (3) hold. This yields the expected utility $EW_{\text{never}} = -(1-q)(b_1 - b_2)^2 - (1-q)R$ which is only higher than its expected utility from always sanctioning $EW_{\text{always}} = -qS - (1-q)(b_1 - b_2)^2$ if $qS > (1-q)R$. However, since we consider the case in which (3) holds by assumption, this can never be true. Hence, (3) implies that $qS < R(1-q)\frac{\alpha}{1-\alpha}$, which means that $qS > (1-q)R$ cannot hold since $\frac{\alpha}{1-\alpha}$ is always positive.

Now suppose the Commission only sanctions when the received message is b_2. Then, it would obtain the utility $EW_{\text{message}} = -q\alpha S - (1-q)(b_1 - b_2)^2 - (1-q)\alpha R$. This is only higher than $EW_{\text{always}} = -qS - (1-q)(b_1 - b_2)^2$ if $qS > R(1-q)\frac{\alpha}{(1-\alpha)}$. However, since (3) holds by assumption and rewriting (3) yields $qS < R(1-q)\frac{\alpha}{1-\alpha}$, EW_{message} will never be higher than EW_{always} as long as (3) holds.

Second, consider the case in which neither (2) nor (3) holds which means that the Commission should never sanctions. The disciplined government cannot do better than setting its preferred budget b_1. If it would set the higher budget b_2, it would obtain $V_d = -(b_1 - b_2)^2$ which is always lower than $V_d = 0$.

Also the undisciplined government cannot do better than setting its preferred budget b_2. Setting the lower budget b_1 would yield $V_u = -(b_2 - b_1)^2$ instead of $V_u = 0$, whereas the former is always lower than the latter.

The Commission cannot do better either. Suppose it would always sanction while neither (2) nor (3) holds. It would obtain $EW_{\text{always}} = -qS - (1-q)(b_1 - b_2)^2$ which would be higher than $EW_{\text{never}} = -(1-q)(b_1 - b_2)^2 - (1-q)R$ if $qS < (1-q)R$. However, since (3) does not hold by assumption, which means that $qS \geq R(1-q)\frac{\alpha}{1-\alpha}$, EW_{always} is never higher than EW_{never} as long as (3) does not hold.

Now suppose that the Commission would only sanction after receiving $m = b_2$. Then, its utility will be $EW_{\text{message}} = -q\alpha S - (1-q)(b_1 - b_2)^2 - (1-q)\alpha R$. This is higher than $EW_{\text{never}} = -(1-q)(b_1 - b_2)^2 - (1-q)R$ only if $qS < (1-q)R\frac{1-\alpha}{\alpha}$. However, we assess the case in which neither (2) nor (3) holds, which means by rewriting (2) that $qS \geq (1-q)R\frac{1-\alpha}{\alpha}$. Thus as long as neither (2) nor (3) holds, EW_{never} is higher than EW_{message}.

Third, consider the case in which (2) holds and (3) does not. In this case, the disciplined government has no incentive to set another budget than its preferred b_1. Suppose it sets b_2. Then it would obtain $EV_d = -(b_1 - b_2)^2 - (1-\alpha)C$ instead of $EV_d = -\alpha C$. Since $b_2 > b_1$ and $\alpha < \frac{1}{2}$, the former will always be lower than the latter.

Also the Commission has no incentive to change its proposed sanction strategy. If it would always sanction, it would obtain $EW_{\text{always}} = -qS - (1-q)(b_1 - b_2)^2$ and it would obtain $EW_{\text{message}} = -q\alpha S - (1-q)(b_1 - b_2)^2 - (1-q)\alpha R$ if it only sanctions after receiving message $m = b_2$, its suggested strategy. The former is only higher than the latter if $qS < (1-q)R\frac{\alpha}{1-\alpha}$. However, since we consider the case in which (2) holds and (3) does not, we know by rewriting (3) that $qS \geq R(1-q)\frac{\alpha}{1-\alpha}$. Thus, in this case EW_{always} cannot be higher than EW_{message}.

Moreover, if it would never sanction, it would obtain $EW_{\text{never}} = -(1-q)(b_1 - b_2)^2 - (1-q)R$ instead of $EW_{\text{message}} = -q\alpha S - (1-q)(b_1 - b_2)^2 - (1-q)\alpha R$. As a result, never sanctioning is

only more beneficial if $qS > R(1-q)\frac{1-\alpha}{\alpha}$. However, that (2) must hold means that $qS < R(1-q)\frac{1-\alpha}{\alpha}$, so that EW_{never} cannot be higher than EW_{message} in this case.

Appendix B: Equilibrium derivation

To check whether this is an equilibrium, we infer whether any actor has an incentive to deviate from its proposed strategy. First consider the case in which (4) does not hold, so that both the disciplined and the undisciplined government set their preferred budgets b_1 and b_2 respectively. This situation is equal to the one we already considered before. We found that in this case none of the players has an incentive to deviate from its strategy, unless (2) holds, (3) does not, and (4) holds. Therefore, we only need to infer whether any of the actors has an incentive under these conditions.

Accordingly, consider the case in which (2) holds, (3) does not, and (4) holds. The disciplined government has under these conditions no incentive to set another budget than its preferred b_1. If it would set the higher budget b_2 while (5) holds, it would obtain $-(b_1 - b_2)^2 - C$ instead of $-C$. Since $b_2 > b_1$ by assumption, the disciplined government has no incentive to deviate from setting b_1 in this situation. If it would set the higher budget b_2 while (5) does not hold, it would obtain $-(b_1 - b_2)^2$ instead of 0. Again, since $b_2 > b_1$, deviating from b_1 does not pay off.

Also the undisciplined government has no incentive to deviate from its proposed strategy. We proposed the strategy that it sets b_1 when (2) holds, (3) does not, (4) holds, and (5) does not hold. Suppose that the undisciplined government would set b_2 under these conditions. Then we find ourselves in the set of strategies we proposed initially in which the two government types set different budgets. The Commission will use its sanction conditions (2) and (3), rather than condition (5). Given that (2) holds and (3) does not, the Commission will only sanction when it receives message $m = b_2$. The expected utility of the undisciplined government setting b_2 is then $-(1-\alpha)C$. Given that (4) holds, this expected utility is lower than the expected utility setting b_1 which is $-(b_2 - b_1)^2 - \alpha C$. The undisciplined government has therefore no incentive to deviate from its proposed strategy.

Neither the Commission has an incentive to deviate. Given that (4) holds we proposed the strategy to sanction if (5) also holds, and not to impose sanctions if (5) does not hold. Suppose it would not impose sanctions when both (4) and (5) hold. Then, it would obtain $EW = -(1-q)(b_1 - b_2)^2 - (1-q)R$ instead of $EW = -qS - (1-q)(b_1 - b_2)^2$, whereas the former is only higher than the latter when $S > R\frac{1-q}{q}$. However, since (5) holds, which means that $S < R\frac{1-q}{q}$, the former is never higher than the latter, so the Commission has no incentive to deviate.

Finally, suppose that the Commission would impose sanctions when (4) holds and (5) does not. Then it would obtain $EW = -qS$ rather than $EW = -(1-q)R$. The former is only higher than the latter if $S < R\frac{1-q}{q}$, but this is not the case given that (5) does not which means that $S \geq R\frac{1-q}{q}$. So also in this case, the Commission has no incentive to deviate from its proposed strategy.

Appendix C: The Fiscal Rules Cyclicality Index (FRCI)

To measure the cyclicality of fiscal rules we use the FRCI as introduced by Debrun et al. (2008). The index scores are year and country-specific. We construct the index scores ourselves. First, we assign an index score to each fiscal rule in place. Scoring is based on a scheme we outline below. For an overview of all fiscal rules, we use the European Commission's Fiscal Rules Database capturing all national fiscal rules in EU-28 between 1990 and 2015. Second, we multiply the scores with their government coverage to correct for the relative impact on fiscal performance of the total government. That is, scores are multiplied with the percentage of total government spending their targeted sectors represent. Third, we construct a FRCI score for each country-year by adding the scores of all fiscal rules in place in the country-year concerned. When in one year multiple rules apply to the same government level we give unit weight to the rule having the highest FRSI. The other rules receive weights equal to a half, one-third, one-fourth, etc. based on their respective FRSI. In this we follow the European Commission (2017b) to account for the diminishing impact of multiple rules on the same government level.

The scoring scheme for individual fiscal rules as introduced by Debrun et al. (2008) is specified below:

Expenditure rules

2 is assigned for a rule capping expenditure growth or level (in nominal or real terms)

-2 if the rule is defined in terms of an expenditure to GDP ratio

Budget balance, borrowing, and debt rules

1 if the rule is defined in cyclically adjusted terms or if the period for assessing compliance is a full business cycle (> 5 years)

-1 for rules defined over a medium-term horizon (2-4 years)

-2 for rules with a short time horizon (1 year)

Revenue rules

2 is assigned if the rule ensures that cyclical revenues are used for debt reduction, or favors it (the government has to specify in advance how cyclical revenues will be used)

-2 is assigned to other types of rules

MANAGEMENT, STRATEGY AND INNOVATION (MSI)
Naamsestraat 69 bus 3535
3000 LEUVEN, Belgium
tel. + 32 16 32 67 00
msi@econ.kuleuven.be
https://feb.kuleuven.be/research/MSI/