

How fiscal rules matter for government debt reduction: Theory and evidence

Lasse Aaskoven^{*} & Rasmus Wiese[†]

Version: 11 December 2018

Abstract

Fiscal rules are not created equally with respect to coverage and auxiliary institutions. These factors affect the ability of different types of fiscal rules (national or supranational) to limit the deficit bias and ultimately debt. We investigate which fiscal rules are effective in achieving sustained debt reduction, using data from 20 OECD countries. We rely on an updated and extended version of the IMF's fiscal rules database. To keep preferences for fiscal conservativeness 'fixed', we focus on periods where there is an observed preference for fiscal discipline. Our findings contribute to both the literature on the effects of fiscal rules on fiscal policy aggregates and to the literature on what drives successful fiscal adjustments. Specifically, we find that: 1. The European Union's Stability and Growth Pact may have caused sustained debt reduction, while it is more doubtful if the mere existence of national fiscal rules has an effect. 2. Fiscal rules, both national and supranational, have a larger effect on sustained debt reduction when they are embedded in a stricter national institutional framework. The findings highlight that the presence of time inconsistencies is important in determining the ability of fiscal rules to lower government debt.

Keywords: fiscal rules; Stability and Growth Pact; debt reduction; (successful) fiscal adjustment/consolidation.

^{*} University of Essex, Department of Government, email: lasse.aaskoven@essex.ac.uk

[†] University of Groningen, Department of Economics and Business, email: r.h.t.wiese@rug.nl

1. Introduction

Many Western countries have accumulated excessive government debt in a time where the demographic composition of these countries increasingly is putting pressure on public finances. This raises serious concerns about the fiscal sustainability of these countries.¹ Therefore, the discussion is no longer whether these countries should attempt to get their fiscal house in order, but how. Numerical fiscal rules, which set numerical limits for fiscal policy aggregates such as expenditure, deficits and debt, has been brought forward as a potential solution to achieve more sustainable public finances (Schaechter et al. 2012; Wyplosz 2013). It is therefore important to study which types of fiscal rules are actually capable of reducing government deficits such that it has a lasting reducing effect on the debt-to-GDP ratio. While a growing research agenda assesses the impact of fiscal rules (Heinemann et al. 2018), the effects of fiscal rules on government's ability to reduce government debt during periods with an observed preference for fiscal discipline are under-researched. In this paper, we specifically study whether fiscal rules have a lasting effect on the debt-to-GDP ratio, if in place during periods where fiscal consolidation is taking place.

We study both national and supranational fiscal rules. National fiscal rules are increasingly being implemented across countries (Schaechter et al. 2012). For example, Denmark had a number of different fiscal rules based on administrative procedures and coalition agreements in place since the early 1990s and has, since 2014, subject to several fiscal rules with a statutory basis. Other OECD countries as well as a wider variety of the World's countries are increasingly subject to different types of national numerical fiscal rules (Llédó et al. 2017). The results of this article suggest that in the shorter run (1-3 years) the mere existence of national fiscal rules (i.e. not taking heterogeneity of rules into account) is only non-robustly related to lower government debt.

We also investigate the effect of supra-national fiscal rules, in our case only the European Union's Stability and Growth Pact (SGP), which includes numerical fiscal rules for both government debt and government deficits. Previous evaluations of the SGP have generally considered it ineffective due to the absence of effective sanctioning mechanisms in cases of violation; see Hallerberg et al. (2009, 170-198).² However, our results suggest that the SGP have increased national governments' ability to lower government debt.

Furthermore, we find that fiscal rules, both national and supranational, become more effective at achieving long-term debt reduction if they are embedded in a stronger *national* fiscal institutional framework, which includes greater fiscal rule coverage and the existence of formal enforcement procedures as well as stronger supervisory fiscal councils.

Therefore, this article adds to a growing literature, which covers the effect of fiscal rules on government fiscal policy aggregates; see Heinemann et al. (2018) for a meta-regression analysis. This literature increasingly focuses on the macro-economic aspects of fiscal rules and fiscal policy (e.g. Andrés and Doménech 2006; Sacchi and Salotti 2015; Bergman and Hutchison 2015; Krogstrup and Wälti 2008; Grembi et al. 2016, Asatryan et al. 2018). However, whether and how

¹ See Wyplosz (2013) for an excellent review of the factors leading to structural overspending, i.e. fiscal unsustainability as a result of the deficit bias.

² However, Koehler and König (2015) argue that *ceteris paribus*, the Stability and Growth Pact's fiscal rules have actually lowered government debt in European Union countries.

fiscal rules might affect the success of fiscal consolidations has received less attention. Exceptions include Larch and Turrini (2011) who study fiscal consolidations in EU countries, and Guichard et al. (2007) who study OECD countries. Both articles find that the likelihood of successful fiscal consolidation is higher with stricter fiscal rules. However, Lavigne (2011) finds fiscal rules in developed countries to be associated with less need for fiscal consolidation, but does not find a positive effect on fiscal rules on actually implementing fiscal adjustments. Thus, studies of fiscal rules and fiscal consolidations find somewhat mixed results for whether fiscal rules matter for (successful) fiscal consolidations. Furthermore, the previous literature on fiscal rules often does not consider the strictness of the rules as well as the institutional setting in which the fiscal rules are embedded.³ Thus, by making these distinctions, our first contribution is to investigate whether different types of fiscal rules and institutions are able to reduce government debt.

By studying the strength of the auxiliary institutional framework in which the fiscal rules are embedded we also investigate which mechanisms are likely to drive the effect of fiscal rules on the debt-to-GDP ratio. Theoretically, we argue that fiscal rules may contribute to debt reduction via a benchmark (numerator) effect. That is, fiscal rules may serve as a focal point for national governments to reduce excessive deficits and the resulting debt accumulation. Fiscal rules may also work via a credibility (denominator) effect, especially if rational economic actors are less worried about time inconsistency problems in terms of abandonment and non-compliance with a rule (Kydland and Prescott 1977; Wyplosz 2013). That is, fiscal rules may induce credibility by convincing private actors and markets that future debt accumulation, and thus tax payments are permanently reduced. Thereby triggering increased economic activity effectively reducing the debt-to-GDP ratio. With this focus, and to the best of our knowledge, we are the first to investigate whether credibility effects of fiscal rules play a significant role in sustained government debt reductions.⁴ Our empirical results suggest that rules lacking credibility suffer from time inconsistencies. Thus, credibility plays an important role in determining the effectiveness of fiscal rules. Consequently, our second contribution is to clarify the underlying theoretical mechanisms that drive the effect of fiscal rules on government debt.

Methodologically, we take into account the endogeneity of the existence of fiscal rules to preferences for fiscal discipline by keeping these preferences ‘fixed’. We do so by focusing on periods where there is an observed preference for fiscal discipline, i.e. periods where fiscal consolidation is taking place. To identify these periods we employ a new method based on structural break testing, which takes the volatility of within country government budget balances into account.⁵ Doing so, we directly tackle the causality problem that both debt reduction and fiscal rules might be driven by an underlying unobserved variable, namely voter preferences for fiscal discipline (Heinemann et al. 2018; Krogstrup and Wälti 2008; Poterba 1996). Our approach does not rely on assumptions on constant fiscal preference within countries over time, nor does it rely on

³ Exceptions are Turrini (2011) and Guichard et al. (2007).

⁴ Although the link between credibility and successful fiscal consolidation is well established in the literature on fiscal consolidations, (e.g. Alesina and Ardagna 1998, 2010, 2013; Alesina and Perotti 1997; Ardagna 2004; Perotti 1996; Tavares 2004), the potential relationship between credibility and fiscal rules have received much less attention. Exceptions include: Heinemann et al. (2016) who studies credibility effects among German policy makers as result of the German debt-brake rule. And, Lara and Wolff (2014) who studies the credibility of fiscal rules and its effect on government bond risk premia.

⁵ Therefore, it is less prone to measurement error compared to conventional threshold methods. It also identifies longer sequences of consolidations than previous methods; see Wiese et al. (2018).

survey data that is impossible to obtain across countries over a reasonable time frame. This is our third contribution. Our econometric approach to estimate the effects of fiscal rules on government debt relies on both random-effects estimates taking inter-cluster correlation into account, and quasi-experimental double robust estimation to account for selection effects.

The structure of the article is as follows: In section 2, we explain the theoretical mechanisms that can make fiscal rules cause government debt reduction. In section 3, the data and estimation methods are explained. Section 4 discusses the results, while section 5 investigates the robustness of the results. Section 6 concludes.

2. Theory and hypotheses

We expect fiscal rules to positively affect governments' ability to reduce government debt. Theoretically, we can distinguish between a fiscal benchmark (numerator) effect and credibility (denominator) effect through which fiscal rules can cause a sustained debt reducing effect. Below, we review the theoretical justifications for these two effects.

Fiscal Benchmark

Taking a perspective that purely deals with the government embarking on a fiscal consolidation, fiscal rules might increase the chance of debt reduction. The reason is that governments themselves use fiscal rules as a benchmark or focal point for their fiscal policy. Fiscal rules make it easier for the government to stick with a fiscally conservative policy, because they can serve as clear benchmarks and policy goals for the fiscal consolidation program. Otherwise, the plan may suffer from vagueness and contestation both within the government and within the political system as a whole. This argument is similar to the arguments made by Reuter (2015), which argue that although fiscal rules are often not complied with, they serve as benchmarks for government's fiscal policy and thus tends to tilt fiscal policy towards the rules' target(s). Note, that according to this argument, whether the government formally breaks its fiscal rules or not, is not the important aspect of the fiscal rules' effect on fiscal outcomes. Instead, fiscal rules help the government keep a focus on the fiscal consolidation throughout the consolidation period, which should to cause debt reduction.

Thus according to the argument that fiscal rules serve as benchmarks for government fiscal policy, we expect fiscal rules to cause sustained reduction in the debt-to-GDP ratio, regardless of the type of rule. This leads to our first hypothesis:

H1: Fiscal rules positively affect sustained debt reduction during periods with an observed preference for fiscal discipline.

Credibility

The credibility of fiscal policy is important in determining its debt reducing effect. Tavares (2004) showed that the 'Nixon in China effect', where a fiscal consolidation is more likely to be successful if the government chooses fiscal instruments it should be ideologically opposed to. Governments can increase the credibility of consolidations by choosing the fiscal instrument that tends to harm their

electorates (Cukierman and Tommasi 1998). This signals that the adjustment is necessary and that the government is committed to improving the fiscal balance and hence the change is perceived durable. Alesina and Perotti (1997) argue that the composition of the adjustment is important in inducing credibility effects. Specifically, consolidations that are based on expenditure cuts are expected to be successful because they target elements on the budget that have the strongest tendency to increase, namely government wages and welfare programs. In turn, governments that manage to cut these expenditures will gain credibility by showing commitment.

The credibility argument builds on a non-Keynesian framework where fiscal consolidations may cause an output increasing effect if the consolidation is credible. This is opposed to a standard static Keynesian framework, where decreases in public spending/increase in taxes will lead to an output decrease (Bertola and Drazen 1993).

Applying this argument to fiscal rules, fiscal rules can give fiscal policy credibility and therefore in a non-Keynesian framework lead to output expansion (the denominator of debt/GDP may increase) and thus sustained debt reductions via 3 distinct channels:

1. Because (almost) all public spending eventually has to be financed by current or future taxes a permanent reduction in spending leads to a reduction in the present value of all future tax payments. Thus, a credible durable decrease in government spending, which is more likely under fiscal rules (Asatryan et al. 2018), reduces the net present value of future tax payments. This leads to a private wealth effect that eventually increases GDP.

2. Tax increases may also lead to output expansions. If the deadweight loss of taxation depends on the tax rate in a nonlinear way, and the tax increase smoothens the path of taxes over time, it may cause GDP to rise due to a lower deadweight loss of taxes (Blanchard 1990). As Tavares (2004) points out, it may also solve uncertainty over the course of future tax policy. This may also cause economic expansion if the tax policy change is credible and perceived as permanent. This works through the labour market where the wealth effect is likely to dominate if the tax change is permanent. In situations where the tax increase is perceived as temporary, the substitution effect is likely to dominate, causing a decrease in the labour supply and GDP. This logic can be applied to fiscal rules, where fiscal rules might also decrease uncertainty over future policy. Hence, the credibility effect of fiscal rules can be important for whether output will increase or decrease following a fiscal consolidation.

3. Finally, deficit reductions may raise household wealth through decreases in interest rates, because credible cuts in the deficit can lower interest rates. This happens due to lower expected inflation and lower default risk. Similarly, an argument made by Kelemen and Teo (2014) is that fiscal rules can serve as focal points for market actors and help them coordinate to sanction a fiscally irresponsible government through the bond market, a sanctioning which is harder to achieve without clear fiscal rules as focal points. A government anticipating these market reactions would then be more likely to not break their fiscal rules giving the rule credibility. Applying this argument to fiscal rules, we should expect a government, which implements a fiscal consolidation program using fiscal rules, to be more likely to continue with the program for fear of violating their fiscal rules and trigger a negative bond market reaction.

Credibility is key in triggering output expansions following fiscal consolidations (e.g. Alesina and Ardagna 1998, 2010, 2013). We therefore expect that incumbent governments, parliaments and relevant economic actors take into account the credibility of fiscal rules. Thus, the qualitative nature and institutional anchoring of fiscal rules also matter for the effect of fiscal rules on sustained debt reduction during fiscal consolidations. Specifically, we expect that the auxiliary institutional framework and the legal base of and scope of fiscal rules contribute to credibility of a fiscal rule. For example, a fiscal rule is more credible if it covers a larger part of the public sector, if it has a stronger legal basis and if auxiliary institutions, such as an independent fiscal council that monitors compliance with sanctioning options. Consequently, we expect fiscal rules with stronger auxiliary frameworks to be more effective. This leads to our second hypothesis.

H2: The stronger the institutional framework in which the fiscal rules are embedded, the stronger is their effects on sustained debt reduction.

Additionally, a fiscal rule that is costly to abandon or to not comply with may limit time inconsistency problems. Rational agents will perceive it to be less likely that a government breaks or even cancels such a rule and therefore they will perceive this rule more credible (Kydlan and Prescott 1977). In the case of supra-national fiscal rules, like the SGP, it is obviously very costly to abandon the rule, i.e. to leave the Eurozone. So, we expect that rules that are costly to abandon or break increase the effectiveness of the rule by inducing credibility.

3. Data and identification strategy

To test our hypotheses, we use a dataset of 20 OECD countries in the years from 1967-1989 to 2013.

3.1 Identifying periods with observed preference for fiscal discipline

Since we aim to estimate the causal debt-reducing effect of fiscal rules, we have to consider the identification problem that both fiscal rules and debt reduction are likely to at least partially be driven by an underlying variable, namely preferences for fiscal discipline (Heineman et al. 2018; Krogstrup and Wälti 2008; Poterba 1996). However, such preferences are hard, if not impossible, to measure directly. Therefore, we focus on an outcome-based proxy for these preferences, namely whether the government is consolidating fiscally, i.e. whether a fiscal adjustment is taking place. We focus only on these periods, as we are sure that politicians are keen on fiscal discipline and then analyse if fiscal rules help achieve the objective of reducing the debt-to-GDP ratio.

Our approach to identify the beginning of a period with observed preference for fiscal discipline (i.e. fiscal adjustment) is based on the identification of changes in the Data Generating Process of fiscal variables, for example, as a result of a fiscal adjustment. Bai and Perron (1998, 2003) develop a general method for this purpose. Consider a model with m possible structural breaks⁶:

$$y_t = \delta_j + \mu_t \quad (t=1, \dots, T; j=1, \dots, m+1) \quad (1)$$

⁶ This part of the paper draws on Wiese et al. (2018).

Where y_t is the dependent variable, in our case the cyclically adjusted primary budget balance in each individual country separately, δ_j is a vector of estimated constants, i.e. the mean at the $m + 1$ different segments of the time series y_t and u_t is the error term. The Bai and Perron (BP) filter generates the segmented route through the series that yields the lowest Sum of Squared Residuals (SSR) up to a maximum number of breaks. The maximum number of breaks is restricted by a trimming parameter h , which specifies a minimum number of observations that have to occur between consecutive breaks. We have set $h=0.15$.⁷ The process underlying the algorithm is straightforward. First, it searches for all possible sets of breaks up to a maximum, restricted by the trimming parameter h , and determines for each number of breaks the set that minimizes the SSR. Then, F-tests determine whether the improved fit produced by allowing additional breaks is sufficiently large, compared to what can be expected randomly, on the basis of the asymptotic distribution derived in Bai and Perron (1998). We use the test procedure recommended by Bai and Perron (2003) to select the optimal number and timing of breaks. That is, dependent on properties of the individual time series, we chose the appropriate filter specification and test. Generally though, the error distribution is allowed to differ across segments.⁸ Autocorrelation and potential heteroskedasticity is modelled non-parametrically by running the filter using a Heteroskedasticity and Autocorrelation Consistent estimate of the variance–covariance matrix.

The BP method identifies the break date (fiscal adjustment initiation) as the first year after the structural break. We therefore take a one-year lag to identify the start of the fiscal adjustment. This method will identify the beginning, but not the end of a fiscal adjustment. We decided that the period of fiscal adjustment continues as long as the change in the cyclically adjusted budget balance is positive. We cannot identify breaks in the beginning and end of the sample due to the trimming parameter h (i.e. 0.15 times the country specific sample length). Using this approach, we identify the fiscal adjustments as presented in Table 1. As Table 1 shows, out of 674 yearly observations for the 20 analyzed OECD countries, we identify 108 years with a fiscal adjustment.

The method we use generally identifies adjustments covering multiple years, as opposed to most previous threshold methods that generally identify short-term changes in fiscal variables. See Wiese et al. (2018) for a comparison of the different approaches. Additionally, threshold methods suffer from measurement error because they ignore the differences in variability in fiscal balances between countries.⁹ Threshold methods work after a one-size-fits-all principle. However, fiscal institutions, and hence the variability of fiscal balances differ between countries. The used method takes such differences into account. Ignoring differences in variability lead to a pattern in the identified adjustment periods where countries with volatile budget balances are over-represented and countries with stable budget balances are under-represented. As a consequence, there would be measurement error in the periods we analyse if falling to take account of the variability of fiscal balances across countries.

⁷ The researcher sets the trimming parameter prior to the analysis.

⁸ This means that we do not assume constant fiscal institutions across time.

⁹ This difference is most likely a result of fiscal institutions differing between countries.

Table 1. Periods with observed preference for fiscal discipline

Country and sample length:	Fiscal discipline/adjustment periods (based on Bai-Perron tests; 5% significance level)
Australia, 1989-2013	1996-98
Austria, 1977-2013	1995-97
Belgium, 1971-2013	1983-87, 1992-98
Canada, 1970-2013	1986-89, 1995-97
Denmark, 1973-2013	1984-86, 1998-99
Finland, 1977-2013	1995-98
France, 1978-2013	1995-99, 2001
Germany, 1970-2013	1981-85
Iceland, 1980-2013	1990-92, 1995-97
Italy, 1970-2013,	1981-83, 1991-93
Japan, 1971-2013	1983-90, 2005-06
Netherlands, 1971-2013	1995-97
New Zealand, 1987- 2013	
Norway, 1980-2013	1994-97
Portugal, 1981-2013	1984, 2009-13
Spain, 1979-2013	1986-87, 1995-99
Sweden, 1973-2013	1984, 1996-98
Switzerland, 1990-2013	1998-99, 2005-08
United Kingdom, 1972-2013	1979-82, 1995-00
United States, 1967-2013	1994-98
Total no. of years with an observed preference for fiscal discipline	108

Notes: This table shows the identification of periods with fiscal adjustments based on the Bai-Perron test. Luxembourg, Greece and Ireland were excluded from the analysis because we had too few observations for these countries to run the Bai-Perron filter.

3.2 Fiscal Rules

Our key independent variables are different measures of numerical fiscal rules. We follow Schaechter et al. (2012) and define fiscal rules as rules of any kind,¹⁰ which set some numerical limit on a fiscal policy aggregate such as government expenditure, revenue, deficits and debt. Examples include expenditure ceilings, which cap government spending at a certain level,¹¹ balanced budget rules that prohibit deficits above a certain threshold or prohibit the government from borrowing to fund current expenditures,¹² as well as debt rules which set a maximum level for government debt.¹³ We use two types of variables to measure fiscal rules:

The first are dummies that takes the value 1 if one or more fiscal rules are in place, such as expenditure rules, revenue rules, balanced budget rules and debt rules. We distinguish between four types of fiscal rules dummy variables:

1. Any fiscal rule, national or supranational, in place.
2. A national fiscal rule in place.
3. A supranational fiscal rule in place.
4. A supranational fiscal rule and a national fiscal rule in place at the same time.

This allows us to test the independent effect of national and supra-national fiscal rules, but also whether there is a joint effect. For OECD countries during the analysed period, the only supranational fiscal rule is the European Union's fiscal rules, from 1999 known as the SGP.¹⁴ The other type of variable we use is indexes, which measure both the existence of each type of fiscal rule (national or supranational), but importantly also their legal scope, their coverage and whether official enforcement mechanisms and auxiliary institutions and rules exists for the rules. These indexes are used to test hypothesis 2. We use three different types of fiscal rule strength indexes:

1. Captures the strength of the national fiscal rules framework.
2. Captures the strength of the supranational fiscal rules framework (de-facto the strength of the SGP).
3. Captures the strength of SGP including national support institutions. This variable measures both the strength of the supranational fiscal rule framework, but takes into account auxiliary national fiscal institutions and rules supporting the SGP.

We describe the construction of these fiscal rules strength indexes in detail in appendix A.

It would have been interesting to analyse if there are heterogeneous effects of different types of national fiscal rules, such as balanced budget rules, debt rules, expenditure rules and revenue rules. However, the high correlation between those rules causes severe multicollinearity in the regressions when controlling for different rule types, e.g. high correlation between expenditure and balanced budget rules. At the same time, we do not trust estimates of the effect of specific national rule types

¹⁰ The range from publically stated government priorities to being set in a country's constitution.

¹¹ In place in Sweden since 1997.

¹² Such a provision was part of the German constitution 1969-2010.

¹³ 60 % of GDP for countries under the European Union's Stability and Growth Pact and later Fiscal Compact.

¹⁴ The Stability and Growth Pact contained both a debt rule, which prohibited government debt over 60 % of national GDP and a balanced budget rule, which prohibited fiscal deficits over 3 % of national GDP

without controlling for other rules, exactly because of this high correlation. In other words, we cannot identify the effect of any specific rule type.

The main source for the existence and characteristics of fiscal rules in the International Monetary Fund's (IMF) fiscal rules' database, which contain information about fiscal rules for all countries, which had any of these in place from 1985 and onwards. The IMF's fiscal rules' dataset have been extended back to 1967 using both information from the database's supplementary material (Bova et al. 2015; Lledó et al. 2017) and independent searches and background research.¹⁵ An overview of the different types of fiscal rules in place in the countries in our sample can be found in table 2.

¹⁵ Specific sources are available upon request.

Table 2. Fiscal rules in sample countries

Country, time period in dataset	National fiscal rule(s) in place	Supranational fiscal rule(s) in place
Australia, 1989-2013	Expenditure rule (1985-1988, 2009-) Revenue rule (1985-1988, 1998-) Balanced budget rule (1985-1988, 1998-) Debt rule (1998-)	None.
Austria, 1977-2013	Balanced budget rule (1999-)	Balanced budget rule (1995-) Debt rule (1995-)
Belgium, 1971-2013	Expenditure rule (1993-1998) Revenue rule (1992-1999) Balanced budget rule (2013-)	Balanced budget rule (1992-) Debt rule (1992-)
Canada, 1970-2013	Expenditure rule (1998-2005) Balanced budget rule (1998-2005) Debt rule (1998-2005)	None
Denmark, 1973-2013	Expenditure rule (1994-2007, 2009-2014) Revenue rule (2001-2011) Balanced budget rule (1992-)	Balanced budget rule (1992-) Debt rule (1992-)
Finland, 1977-2013	Expenditure rule (2003-) Balanced budget rule (1999-) Debt rule (1995-2006)	Balanced budget rule (1995-) Debt rule (1995-)
France, 1978-2013	Expenditure rule (1998-) Revenue rule (2006-) Balanced budget rule (2012-)	Balanced budget rule (1992-) Debt rule (1992-)
Germany, 1970-2013	Expenditure rule (1982-2009) Balanced budget rule (1969-)	Balanced budget rule (1992-) Debt rule (1992-)
Iceland, 1980-2013	Expenditure rule (2004-2008)	None.
Italy, 1970-2013,	None.	Balanced budget rule (1992-) Debt rule (1992-)
Japan, 1971-2013	Balanced budget rule (1947-) Expenditure rule (2006-2008, 2010-2012)	None.
Netherlands, 1971-2013	Expenditure rule (1994-) Revenue rule (1975-1979, 1994-) Balanced budget rule (1961-1974, 1980-1982, 1994-)	Balanced budget rule (1992-) Debt rule (1992-)
New Zealand, 1987- 2013	Balanced budget rule (1994-) Debt rule (1994-)	None.
Norway, 1980-2013	Balanced budget rule (2001-)	None.
Portugal, 1981-2013	None.	Balanced budget rule (1992-) Debt rule (1992-)
Spain, 1979-2013	Expenditure rule (2011-) Balanced budget rule (2003-2011)	Balanced budget rule (1992-) Debt rule (1992-)
Sweden, 1973-2013	Expenditure rule (1997-) Balanced budget rule (2000-)	Balanced budget rule (1995-) Debt rule (1995-)
Switzerland, 1990-2013	Balanced budget rule (2003-)	None.
United Kingdom, 1972-2013	Balanced budget rule (1997-) Debt rule (1998-2008, 2010-)	Balanced budget rule (1992-) Debt rule (1992-)
United States, 1967-2013	Expenditure rule (1990-2002, 2011-) Balanced budget rule (1986-1989)	None.

Note: Some of the countries shifted to another version of the same type of fiscal rule during the period. Data is based on Lledó et al. (2017) and own coding.

3.3 Control variables and estimation approach

We focus on the effect of fiscal rule in terms of its impact on government debt. Therefore, we start by defining the factors that make up the debt-to-GDP ratio in year t , $(D/Y)_t$:

$$(D/Y)_t \equiv \left(\frac{D}{Y}\right)_{t-1} + \left(\frac{Bal}{Y}\right)_t + (iD_{t-1}/Y)_t \quad (2)$$

Where D is the debt, Y is GDP, $Bal (=T-G)$ is the primary balance, i.e. taxes (T) minus expenditures (G) both excluding interest payments, and iD are the interest payments on D . So, the debt-to-GDP ratio in year t equals the amount of outstanding debt from the previous period, plus the deficit in the current period, plus interest payments on outstanding debt. If we move the first term on the right-hand side of eq. (2) to the left side, add $\left(\frac{Bal}{Y}\right)_{t-1}$ on both sides and simplify we obtain eq. (3), where Bal has been substituted with $(T-G)$ in the right most expression:

$$\Delta\left(\frac{D}{Y}\right) = \Delta\left(\frac{Bal}{Y}\right) + \left(\frac{Bal}{Y}\right)_{t-1} + (iD_{t-1}/Y)_t = \Delta((T - G)/Y) + \left(\frac{Bal}{Y}\right)_{t-1} + (iD_{t-1}/Y)_t \quad (3)$$

Alesina and Ardagna (2010), amongst others, concentrate on the first component on the right hand side of this equation, namely the fiscal impulse when estimating what causes debt reduction. Thereby, they ignore the initial balance and the interest payments on outstanding debt, which is driven by the interest rate and the amount of outstanding debt. In our empirical specification, we therefore, in addition to the fiscal impulse, include the lagged cyclically adjusted budget balance, the lagged debt-to-GDP ratio, and the short-term interest rate as well real GDP growth (cf. Baldaci et al., 2004; 2012). These variables capture the need for fiscal discipline within periods of fiscal adjustment, thus they are likely to correlate both with the implementation of fiscal rules and the debt-to-GDP ratio. Also, initial conditions are important in determining the credibility of fiscal rules and therefore also whether they are effective in reducing debt (Heinemann et al. 2016).

According to Alesina and Perotti (1997), Alesina and Ardagna (2010) and Perotti (1996), an important trigger of whether governments succeed in lowering the debt-to-GDP ratio is the choice of fiscal instrument during adjustment periods. This might be correlated with the type fiscal rule in place in a given country. Therefore, the change in expenditure and revenue are added as controls. Additionally, we control for welfare state size, since this might matter for the effect of fiscal rules on fiscal outcomes, confer Heineman et al. (2018). Lastly, countries may use “creative accounting” to display adherence to a rule, especially when it becomes binding (Wyplosz 2013; Alesina and Perotti 1996). So, countries may have a rule in place, but instead of behaving according to the rule, which potentially could reduce debt, they start misreporting national account figures and continue to accumulate debt. Therefore, we include the measure proposed by Alt et al. (2014) for stock-flow adjustments (SFA) in national accounts. In year t , it is defined as $SFA_t = D_t - D_{t-1} + B_t$, where the first two terms on the right hand side are the annual change in gross debt (D) and B is the non-cyclically adjusted budget balance, in which a deficit is expressed as a negative number. An increasing deficit would, on average increase debt, thus a positive SFA indicates that the change in gross debt exceeds the magnitude of the budget deficit (or decreases by less than a surplus). For example, if outstanding debt increases by 4 and the balance is reported as 2 (a deficit), then the

above expression yields an SFA of 2. A surplus of 2 that resulted in no debt reduction gives the same result. A negative SFA implies that government debt increases by less than the size of the deficit. Appendix B provides sources and summary statistics of all variables and a correlation matrix.

Our estimation strategy follows two approaches. First, we estimate random effects models of the effect on the 3- and 5-year change in the debt-to-GDP ratio.¹⁶

$$y_{it} = \beta_1 rule_{it} + \beta_x x'_{it} + \varepsilon_{it} \quad (4)$$

Where y_{it} is the change in debt after 3 or 5 years at a given country-year observation, $rule_{it}$ is either a dummy variable indicating whether a given type of rule is in place, or a continuous variable measuring the strength of a rule if it is in place, x'_{it} is a vector of the control variables discussed above and ε_{it} is the error component.

Secondly, in the robustness analysis section, we use augmented inverse probability weighted regressions to account for potential selection bias. Countries that adopt fiscal rules may be the ones that expect to benefit from it in terms of debt reduction. The augmented inverse probability weighted regression consists of two steps. First, the likelihood of having a fiscal rule (of any type) in place is estimated in a latent variable framework. Consider a panel discrete choice model in which the dependent variable is a dummy equal to 1 if there is a fiscal rule in place in country i at time t ($y_{i,t}$) and 0 when there is not. The idea is that countries will adopt a fiscal rule if their expected benefit from doing so is positive:

$$\begin{aligned} y_{it} &= 1 & \text{if } y_{it}^* > 0 \\ y_{it} &= 0 & \text{if } y_{it}^* \leq 0 \end{aligned} \quad (5)$$

Where: $y_{it}^* = x'_{it}\beta + v_{it}$. This function can be interpreted as the likelihood that a country will benefit and therefore has a rule in place, which is dependent on observed variables (x), and a random error term (v). We use the same variables as explained in equation 3 plus the size of government and the inflation rate to estimate the probability that a country has a fiscal rule in place in a given year. These variables are expected to drive the need for debt reduction and hence determine the government's inclination to adopt a fiscal rule. The probability that we observe that a country has a fiscal rule in place is:

$$P(y_{it} = 1) = P(y_{it}^* > 0) = P(v_{it} > -x'_{it}\beta) = F(x'_{it}\beta) \quad (6)$$

So, it depends on the countries (expected) benefit from implementing a rule. F is the logistic probability function that ensures that the estimated probability of observing a fiscal rule in place is bounded between zero and one.

Secondly, the estimated probabilities of whether a country has a rule in place are used as weights in a weighted least squares regression, specified as equation (4). Specifically, observations where a rule is in place get a weight by the inverse of the probability score, whereas the observations where no rule is in place get a weight of the inverse of one minus the probability score. This means that

¹⁶ We believe that a 3- and 5-year time windows are relevant evaluation periods. See for example the literature evaluating the effects of fiscal adjustments on government debt.

observations with a rule in place and a low probability score get a higher weight in the regression, along with observations with no rule in place and a high probability score. This puts more weight on observations that are comparable and hence reduces selection bias. The augmented weighting adds an adjustment factor to the treatment effect when the estimated probability scores are close to zero or one. The method is said to be double robust, as it only requires one of the following two conditions to hold: The conditional mean model is correctly specified, or the probability score model is correctly specified. Weighting can be interpreted as removing the correlation between the covariates and the rule indicator, and regression removes the direct effect of the covariates, see Robins and Rotnitzky (1995), Robins et al. (1995), and van der Laan and Robins (2003) for more detail. We report the Average Treatment Effect (ATE), which is calculated as the average difference between countries with a rule in place and countries without a rule based on the weighted OLS regression line for both groups.¹⁷

4. Results

Before we continue with the regression analysis, we present some simple descriptive evidence of the relationship between fiscal rules and debt accumulation with the aim of explaining the nature of the data.

4.1 Descriptive evidence

To gauge the relationship between fiscal rules and the debt-to-GDP ratio, we begin by calculating the average change in the debt-to-GDP ratio over the 3 and 5 years after a rule was in place in a given country-year combination. Comparing the averages for any type of rule clearly shows that fiscal rules are associated with debt reduction. We only analyse periods of fiscal consolidation, thereby ensuring a significant degree of preferences for fiscal discipline. However, preferences for fiscal discipline may be stronger in countries with for example large deficits or high debt-to-GDP ratios. Therefore there it is necessary to include the control variables discussed above.

Tables 3 and 4 also suggest that the debt reducing effect is strongest when both a national and supranational fiscal rule is in place simultaneously. This suggests a complementarity effect. However, the simple average comparison ignores the effect that may be caused by either type of rule on its own. This highlights the need to control for those effects when isolating a potential complementarity effect, as is done in the regression analysis.

¹⁷ Although it would be interesting to take account of selection effects concerning rule strength the inverse probability weighted regression framework is only suitable for a binary indicator of treatment.

Table 3: 3-year average change in debt/GDP after years with a fiscal rule in place vs. no fiscal rule

	Any fiscal rule in place	Any national fiscal rule in place	SGP	National rule & SGP joint
Yes	-3.679 (10.274) n=67	-5.563 (8.645) n=50	-4.910 (10.993) n=42	-9.517 (6.579) n=25
No	0.180 (10.004) n=41	0.674 (10.795) n= 58	-0.497 (9.522) n=66	-0.014 (10.226) n=83
T-test for difference in means, p-value	0.029	0.001	0.018	0.000

Note: The table shows average debt-to-GDP ratio in years with a rule in place vs. years with no rule in place, in years with an observed preference for fiscal discipline. Standard deviation of the average is shown in parentheses. The t-test for difference in the mean of debt after 3 years is based on a one-sided (H_0 : Yes>no) non-paired hypothesis test. National rule & SGP joint is dummy equal to one if there is both a national and supranational rule in place at the same time, otherwise the variable is equal to zero.

Table 4: 5-year average change in debt/GDP after years with a fiscal rule in place vs. no fiscal rule

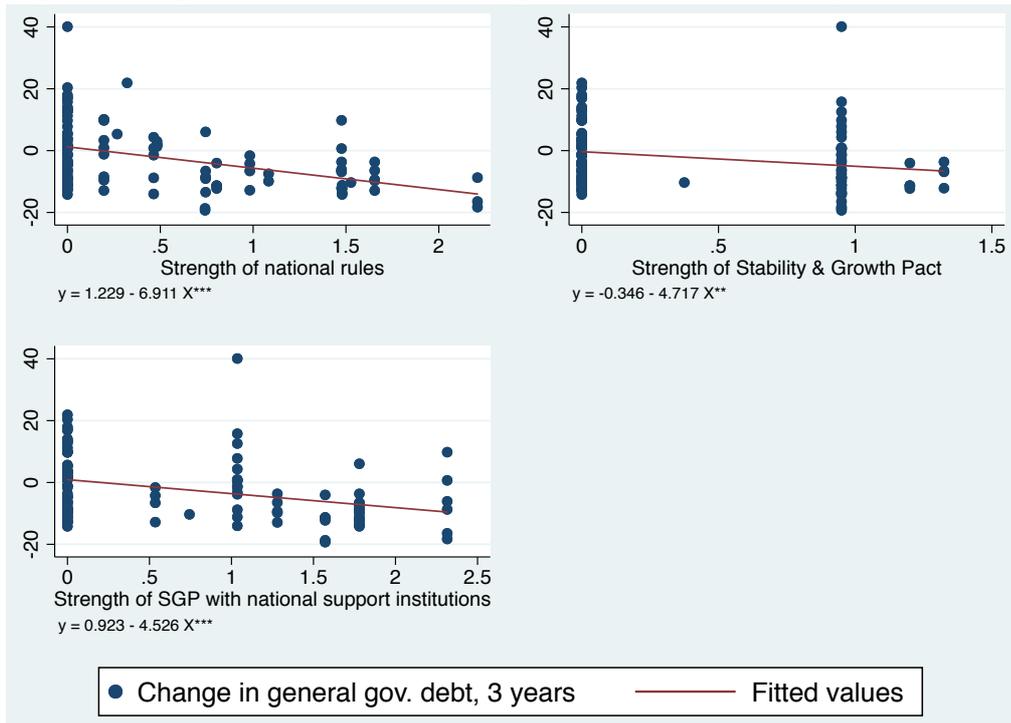
	Any fiscal rule in place	Any national fiscal rule in place	SGP	National rule & SGP joint
Yes	-6.115 (13.296) n=66	-6.911 (13.549) n=50	-9.718 (11.879) n=41	-13.616 (9.801) n=25
No	1.812 (13.681) n=41	0.286 (13.491) n= 57	1.048 (13.583) n=66	0.136 (13.443) n=82
T-test for difference in means, p-value	0.002	0.004	0.000	0.000

Note: The table shows average debt-to-GDP ratio in years with a rule in place vs. years with no rule in place, in years with an observed preference for fiscal discipline. Standard deviation of the average is shown in parentheses. The t-test for difference in the mean of debt after 5 years is based on a one-sided (H_0 : Yes>no) non-paired hypothesis test.

Fig. 1 and fig. 2 plots the change in debt over 3 and 5 years after a given country-year combination on the y-axis against the three types of fiscal rule strength indexes used in our analysis. Years with a fiscal rule strength value equal to 0 represent years where no rule was in place. The figures show a decreasing effect on debt as the strength of a given index increases. The figures also show that modelling this linearly with a continuous variable in the regression analysis is a reasonable approximation of its effect on debt.

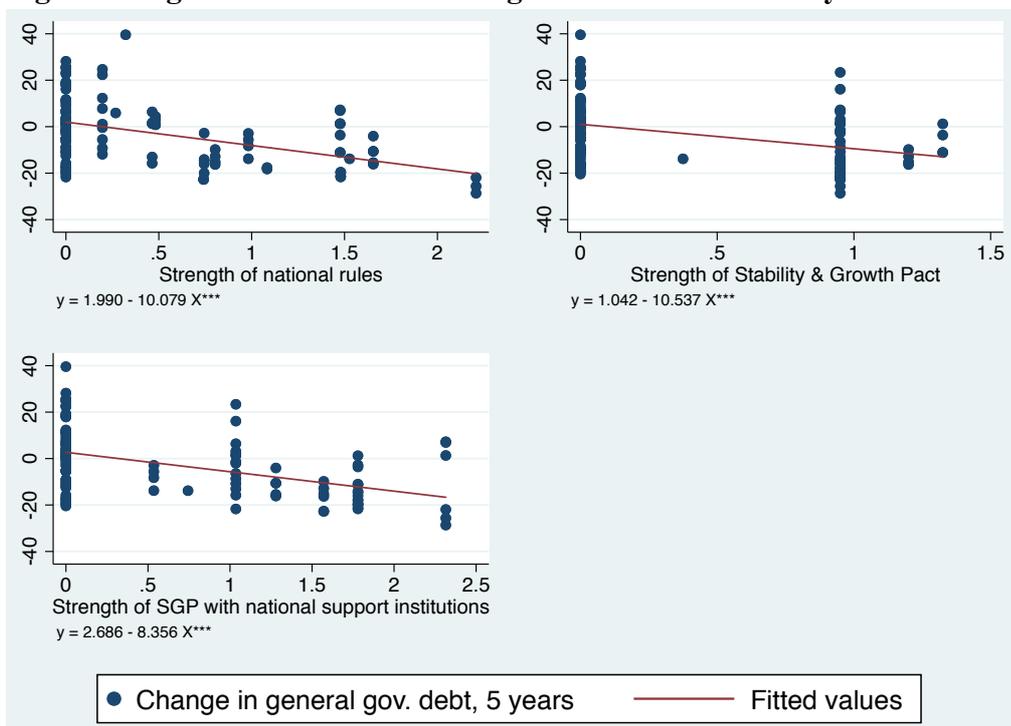
Overall, the descriptive statistics leads us to expect an effect of fiscal rules on the debt-to-GDP ratio. However, to investigate whether there is a causal effect of rules on debt we need to control for factors that affect both the adoption of fiscal rules and the need for debt reduction. This is done in the regression analysis below.

Fig. 1. Strength of rules and the change in debt/GDP after 3 years



Note: Significance of the slope of the bivariate regression line is indicated in the following way: *** $p < 0.01$, ** $p < 0.05$

Fig. 2 Strength of rules and the change in debt/GDP after 5 years



Note: Significance of the slope of the bivariate regression line is indicated in the following way: *** $p < 0.01$, ** $p < 0.05$

4.2 Regression results

In table 5 and 6, we report the regression results of a random effect model specification. The Wald-test of the Mundlak (1978) averages shows that a random effects model is appropriate. In table 5.1 and 6.1, we calculate bootstrapped fiscal rules coefficients along with the wild-cluster bootstrapped standard errors of the same models based on the method suggested in Cameron et al. (2008) due to the problem explained in Bertrand et al. (2004).¹⁸ Table 5 and 5.1 show the estimated effect after three years, and table 6 and 6.1 show the estimated effect after five years.

Focusing on the effect of fiscal rules on the 3-year change in the debt-to-GDP ratio, model 1 tests the effect of the dummy variable ‘*any rule in place*’. The results show that a broad measure of fiscal rules that does not distinguish between type(s) of rules in place does not significantly affect debt reduction in a period of three years. Model 2 – 5 includes dummies for SGP and national fiscal rules, first separately, then together, and lastly with an interaction term to capture a potential complementarity effect. Given the results, when taking account of within cluster correlation in table 5.1, we conclude that the complementarity effect is insignificant and therefore model 4 is the most appropriate specification. Hence, the marginal insignificant coefficient of ‘*any national fiscal rule*’ in model 4 in table 5.1 suggests that over three years, the existence of a ‘*national fiscal rule*’ reduces the debt-to-GDP ratio by around 4.5 percentage points over 3 years.

None of the rule strength indicators are significant when using the wild bootstrap method. Hence, we conclude that over a three-year evaluation period, the institutional and legal frameworks supporting the rules do not play a significant role in reducing debt. Therefore, the short-term debt reducing effect of national fiscal rules seems to mostly be caused by a benchmark effect and not a credibility effect. We do note though, that both ‘*strength of national rules*’ and ‘*strength of SGP with national support institutions*’ is approaching statistical significance in table 5.1.

¹⁸ The problem is that rules are in place as sequences over time within countries, see table 2. Like difference-in-difference estimation this leads to autocorrelation within countries. Hence, we need to calculate clustered standard errors using our unbalanced panel with few clusters.

Table 5: Random effects estimates on debt/GDP after 3 years

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(6)	(8)
	debtgdp3							
Change in expenditures	1.991 (0.115)	2.034* (0.096)	2.107* (0.094)	2.168* (0.081)	1.968 (0.109)	2.083* (0.079)	2.071* (0.099)	2.321* (0.056)
Change in revenues	-3.483*** (0.010)	-3.677*** (0.005)	-3.594*** (0.007)	-3.900*** (0.003)	-3.684*** (0.005)	-3.981*** (0.002)	-3.556*** (0.007)	-3.940*** (0.002)
Adjustment size	-0.034 (0.964)	0.070 (0.925)	0.003 (0.997)	0.161 (0.829)	0.114 (0.876)	0.144 (0.844)	0.006 (0.994)	0.074 (0.920)
Lagged CAPB	-1.276*** (0.000)	-1.108*** (0.002)	-1.222*** (0.001)	-1.018*** (0.007)	-0.925** (0.015)	-0.968*** (0.005)	-1.225*** (0.001)	-1.089*** (0.002)
Lagged debt/GDP	0.009 (0.806)	0.011 (0.758)	0.012 (0.724)	0.010 (0.774)	0.015 (0.675)	0.025 (0.382)	0.010 (0.784)	0.025 (0.417)
GDP growth rate	-1.515** (0.022)	-1.453** (0.026)	-1.502** (0.023)	-1.460** (0.026)	-1.385** (0.034)	-1.544** (0.012)	-1.477** (0.025)	-1.455** (0.021)
Interest rate	0.092 (0.728)	0.031 (0.893)	0.094 (0.689)	-0.033 (0.895)	0.132 (0.619)	-0.050 (0.813)	0.104 (0.660)	0.006 (0.980)
Size of welfare state	-0.152 (0.247)	-0.175 (0.182)	-0.077 (0.602)	-0.110 (0.476)	-0.017 (0.919)	-0.180* (0.089)	-0.085 (0.565)	-0.077 (0.523)
Stock-flow adjustments	0.144 (0.539)	0.082 (0.724)	0.157 (0.502)	0.085 (0.714)	0.031 (0.894)	0.003 (0.988)	0.155 (0.507)	0.124 (0.590)
Any rule in place	-1.511 (0.520)							
Any national fiscal rule		-4.012* (0.057)		-3.898* (0.072)	0.048 (0.988)			
Stability & Growth pact			-2.668 (0.263)	-1.957 (0.424)	1.247 (0.679)			
National rule & SGP joint					-7.621* (0.075)			
Strength of national rules						-4.889*** (0.001)		
Strength of SGP							-2.364 (0.304)	
Strength of SGP with national support institutions								-3.168** (0.013)
Constant	9.023 (0.209)	11.020 (0.115)	5.972 (0.385)	9.765 (0.187)	3.016 (0.719)	12.076** (0.036)	6.189 (0.370)	7.114 (0.232)
Observations	108	108	108	108	108	108	108	108
Number of identifier	19	19	19	19	19	19	19	19
R-squared overall	0.417	0.426	0.426	0.433	0.456	0.475	0.424	0.456
Wald-test, Mundlak averages	0.026	0.052	0.032	0.542	0.326	0.072	0.034	0.135

Note: P-values in parentheses. *: p<0.10, **: p<0.05, ***: p<0.01.

Table 5.1: Random effects estimates of the effect of rules on debt/GDP after 3 years, wild cluster bootstrap coefficients and standard errors

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimate	-1.784	-4.845*	-2.957	-4.355	0.363	-4.259	-2.374	-2.918
(p-value)	(0.392)	(0.088)	(0.331)	(0.107)	(0.868)	(0.175)	(0.474)	(0.167)
				-1.235 (0.668)	1.481 (0.584)			
					-7.587 (0.173)			

Note: The wild cluster bootstrap is based on 1000 replications where re-sampling takes place at the country level. The P-values are calculated using a conservative number of degrees of freedom, namely 18, which equals the number of clusters minus 1.

In table 6, we estimate the effect of fiscal rules on the 5-year change in the debt-to-GDP ratio. Model 1 of table 6 and 6.1 indicates that the catch-all variable '*any rule in place*' is significant when taking account of within cluster correlation. Model 2 - 5 of table 6 and 6.1 leads one to suspect that this benchmark effect is mainly driven by the '*Stability & Growth Pact*' and that the SGP has been an effective debt-reducing tool for the countries joining the union. For national rules, there is not much evidence in favour of a benchmark effect. However, the '*strength of national rules*' indicator in table 6.1 suggests a debt reducing effect of 5.39 percentage points of GDP for every 1-unit increase in national fiscal rule strength, although the variable is marginally insignificant. Based on table 6.1, we can conclude that the catch-all variable '*any rule in place*' is significant along with the '*strength indicator of SGP with national support institutions*'. This indicates that both a benchmark effect and a credibility effect may be behind the debt reducing effect of the SGP variables.

In conclusion, in the shorter-run, national fiscal rules seem to be mainly effective in producing debt reduction through a benchmark effect. In the longer run, the results show a significant debt-reducing effects of having '*any rule in place*', mainly driven by the SGP. Also the '*strength of SGP with national support institutions*' is significant. The results indicate that for longer-term government debt reduction, supranational fiscal rules (in our case the SGP) can actually be effective. However, their effectiveness also depends on whether the supranational fiscal rules are supported by a strong national fiscal institutional framework. This suggests that the SGP has impacted government debt through a credibility mechanism. Although the (marginally) insignificant effect of the SGP dummies in table 6.1 also indicates that the SGP might also have independently reduced government debt through a benchmark effect. Note also that the '*strength of national rules*' is marginally insignificant in table 6.1 and that the '*strength of SGP*' at the European level does not matter. These findings carry over from table 5.1. Hence we have consistent evidence pointing in the direction that the strength of national auxiliary institutions supporting rules, whether national fiscal rules or supranational fiscal rules matter. This indicates that national institutions may help rules gain credibility.

Table 6: Random effects estimates on debt/GDP after five years

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	debtgdp5	debtgdp5	debtgdp5	debtgdp5	debtgdp5	debtgdp5	debtgdp5	debtgdp5
Change in expenditures	0.687 (0.718)	0.567 (0.764)	0.804 (0.667)	0.923 (0.623)	0.861 (0.648)	0.859 (0.644)	0.678 (0.718)	1.083 (0.558)
Change in revenues	-1.902 (0.334)	-1.714 (0.376)	-2.086 (0.280)	-2.328 (0.232)	-2.241 (0.252)	-2.158 (0.261)	-1.913 (0.324)	-2.406 (0.209)
Adjustment size	-0.305 (0.767)	-0.369 (0.718)	-0.218 (0.830)	-0.139 (0.891)	-0.157 (0.878)	-0.275 (0.790)	-0.242 (0.814)	-0.216 (0.832)
Lagged CAPB	-1.626*** (0.003)	-1.596*** (0.003)	-1.493*** (0.006)	-1.372** (0.014)	-1.326** (0.019)	-1.442*** (0.006)	-1.538*** (0.005)	-1.391*** (0.008)
Lagged debt/GDP	0.069 (0.239)	0.072 (0.223)	0.077 (0.190)	0.077 (0.199)	0.081 (0.180)	0.079 (0.126)	0.072 (0.220)	0.086 (0.103)
GDP growth rate	-0.483 (0.609)	-0.377 (0.690)	-0.493 (0.598)	-0.437 (0.642)	-0.389 (0.680)	-0.598 (0.516)	-0.409 (0.663)	-0.542 (0.554)
Interest rate	0.417 (0.338)	0.555 (0.158)	0.418 (0.292)	0.344 (0.409)	0.439 (0.319)	0.301 (0.401)	0.476 (0.231)	0.268 (0.464)
Size of welfare state	-0.281 (0.236)	-0.309 (0.198)	-0.109 (0.667)	-0.134 (0.614)	-0.082 (0.768)	-0.353* (0.071)	-0.147 (0.563)	-0.181 (0.386)
Stock-flow adjustment	0.061 (0.852)	-0.018 (0.955)	0.095 (0.768)	0.047 (0.884)	0.022 (0.946)	-0.121 (0.715)	0.090 (0.780)	0.019 (0.952)
Any rule in place	-4.671 (0.191)							
Any national fiscal rule		-4.161 (0.205)		-2.784 (0.411)	-0.174 (0.973)			
Stability & Growth pact			-7.045* (0.054)	-6.257 (0.103)	-4.528 (0.325)			
National rule & SGP joint					-4.512 (0.494)			
Strength of national rules						-6.338*** (0.008)		
Strength of SGP							-5.681 (0.110)	
Strength of SGP with national support institutions								-5.754*** (0.005)
Constant	4.055 (0.745)	2.676 (0.828)	-3.237 (0.783)	-0.842 (0.947)	-4.817 (0.731)	8.438 (0.411)	-2.733 (0.818)	2.055 (0.841)
Observations	107	107	107	107	107	107	107	107
Number of identifier	19	19	19	19	19	19	19	19
R-squared overall	0.290	0.287	0.325	0.324	0.338	0.370	0.318	0.375
Wald-test, Mundlak averages	0.168	0.231	0.314	0.542	0.666	0.271	0.329	0.305

Note: P-values in parentheses. *: p<0.10, **: p<0.05, ***: p<0.01.

Table 6.1: Random effects estimates of the effect of rules on debt/GDP after five years, wild cluster bootstrap coefficients and standard errors

Model:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimate	-4.971*	-4.257	-7.319	-2.756	-4.77	-5.390	-5.891	-5.468*
(p-value)	(0.067)	(0.163)	(0.102)	(0.358)	(0.798)	(0.148)	(0.216)	(0.084)
				-6.044 (0.194)	-4.822 (0.224)			
					-3.881 (0.508)			

Note: The wild cluster bootstrap is based on 1000 replications where re-sampling takes place at the country level. The P-values are calculated using a conservative number of degrees of freedom, namely 18, which equals the number of clusters minus 1.

5. Robustness analysis

5.1 Alternative estimation strategies

In this section, we investigate the robustness of our results by applying different estimation strategies. First, we focus on the robustness of the benchmark effect by applying the quasi-experimental augmented inverse probability weighted regression method as described in section 3.3. The main motivation for applying this method is to reduce potential selection bias that may impact the main results (Imbens and Wooldridge 2009). We focus on the Average Treatment Effect (ATE) in terms of the change in the debt-to-GDP ratio after 1, 3 and 5 years. This statistic captures the average effect of treatment in the sample based on potential outcomes. The treatment and control groups are weighted by the estimated probability of having a rule in place; see section 3.3 for detail about the weighting. The estimated logit models are shown in columns 2, 4, and 6 in table 7. As can be seen, the logit model is a good predictor of the presence of fiscal rules.

The results of the outcome regressions are in columns 1, 3, and 5. Columns 1 and 3 show that ‘*any rule in place*’ and ‘*any national fiscal rule*’ does not causes debt reduction after 1, 3 or 5 years. The (benchmark) effect of national fiscal rules is insignificant using this estimation approach. Hence, these results cast serious doubt on the effectiveness of national fiscal rules. Specifically, the results in table 7 concerning national fiscal rules suggest that selection into treatment may have been behind the finding in table 5 that the mere existence of national fiscal rules matters for government debt reduction. However, the ‘*SGP*’ did cause debt reduction after 5 years but not after 1 and 3 years. The magnitude of the effect of SGP is 8.5 percentage points of debt reduction. The estimated effects are in line with the effects found in table 5 and 6, both concerning significance and size. We recognize that selection into treatment may also play a role concerning the strength indexes. However, we are unaware of any method that can handle treatment endogeneity when the treatment variable is continuous.

We do not report the joint effect of national fiscal rules and SGP. As shown in the main results section, we need to control for the individual effects of both national and supranational fiscal rules when estimating the joint effect. However, a model with the joint effect cannot be estimated when including those controls. Model 3 and 5 control for the effect of SGP and national fiscal rules respectively, in the outcome regressions.

Table 7: Augmented inverse probability weighted regression analysis

VARIABLES	Any rule in place		Any national rule		Stability & Growth Pact	
	(1) ATE	(2) Logit	(3) ATE	(4) Logit	(5) ATE	(6) Logit
1 year change in debt	0.215 (0.823)		-0.899 (0.304)		0.705 (0.351)	
3 year change in debt	2.143 (0.411)		-2.428 (0.345)		-2.192 (0.261)	
5 year change in debt	0.544 (0.883)		-1.084 (0.699)		-8.467*** (0.007)	
Lagged CAPB		0.250*** (0.001)		0.338*** (0.000)		0.259*** (0.003)
Lagged Debt/GDP		-0.003 (0.770)		0.001 (0.940)		0.002 (0.777)
GDP Growth		-0.194 (0.251)		-0.011 (0.945)		-0.186 (0.152)
Interest rate		-0.317* (0.074)		-0.264* (0.071)		-0.018 (0.914)
Size of Government		0.002 (0.934)		-0.029 (0.283)		0.187*** (0.000)
Inflation rate		-0.122 (0.490)		-0.145 (0.504)		-0.459* (0.091)
Constant		3.934* (0.087)		3.155* (0.090)		-6.060*** (0.000)
Pseudo R2		0.3368		0.3042		0.4052
Correctly classified		80.56%		75.93%		80.56%
Observations	108	108	108	108	108	108

Note: The regressions for the 5-year change in debt only include 107 observations. In model 3 we control for the ‘SGP’ and in model 5 ‘National fiscal rules’ in the outcome regressions. We do not report the estimates of a joint SGP and national rule effect as the model cannot be estimated with SGP and national rules as controls. Columns 2, 4, and 6 shows the results of the logit model used to estimate probability with which the observations are inversely weighted in the outcome regression. P-values in parentheses: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

In table 8, we take a different approach to measuring the outcome variable. Specifically, we estimate if the presence of fiscal rules is a good predictor of the likelihood of observing a successful fiscal adjustment. Like our identification of periods with fiscal adjustment, on which we base all estimates, we do not rely on conventional threshold methods to classify whether an adjustment is successful or not. Rather, we apply on the Bai and Perron (1998, 2003) filter to identify structural breaks in the growth rate of the debt-to-GDP ratio. Specifically, regimes with negative debt-to-GDP growth rates are compared to the periods with fiscal adjustment. If fiscal adjustments are identified prior to, or simultaneously, with the beginning of regimes with negative debt-to-GDP growth rates, and the periods are not more than 4 years apart, we code it as a successful fiscal adjustment, as this is evidence that the fiscal adjustment had a statistically significant and lasting negative effect on the debt-to-GDP ratio. See Wiese et al. (2018) for more detail and a comparison with conventional threshold methods. Table B4 in the appendix summarizes the results of this procedure to identify successful fiscal adjustments. In total 58 of the 108 periods of fiscal adjustments are classified as successful.

Table 8 reports the marginal effects of a random effects logit specification of the likelihood of success. The results confirm the analysis above. Any rule in place is a significant predictor of

success. The likelihood of observing a negative structural in the growth rate of the debt-to-GDP ratio as a result of ‘*any rule in place*’ increases by 1.7 percentage point, keeping all else equal to the sample means. This result seems mostly to be driven by the SGP, as can be seen in columns 2-5, although the rule coefficients are insignificant in these specifications. Also, the results show a credibility effect of fiscal rules since the ‘*Strength of national rules*’ and ‘*Strength of SGP with national support institutions*’ both are significant. Again it highlights the importance of national auxiliary institutions for the effectiveness of fiscal rules in terms of government debt reduction.

Table 8: Marginal effects of rules on the likelihood of observing a successful fiscal adjustment

VARIABLES	(1) Success	(2) Success	(3) Success	(4) Success	(5) Success	(7) Success	(8) Success	(9) Success
Change in expenditures	-0.406 (0.273)	-0.294 (0.289)	-0.243 (0.339)	-0.386 (0.309)	-0.368 (0.475)	-0.085 (0.755)	-0.242 (0.333)	-0.265 (0.457)
Change in revenues	0.431 (0.260)	0.203 (0.415)	0.220 (0.380)	0.329 (0.352)	0.447 (0.439)	0.010 (0.970)	0.224 (0.365)	0.237 (0.504)
Adjustment size	0.303 (0.166)	0.200 (0.271)	0.223 (0.317)	0.289 (0.236)	0.351* (0.057)	0.301* (0.084)	0.217 (0.310)	0.376** (0.013)
Lagged CAPB	0.175 (0.153)	0.114 (0.154)	0.161 (0.232)	0.178 (0.234)	0.203* (0.092)	0.088 (0.336)	0.156 (0.226)	0.199* (0.077)
Lagged debt/GDP	0.000 (0.971)	-0.003 (0.740)	-0.000 (0.986)	0.001 (0.928)	0.003 (0.744)	0.002 (0.846)	-0.000 (0.977)	0.003 (0.780)
GDP growth rate	0.544** (0.046)	0.358* (0.077)	0.395 (0.101)	0.509 (0.134)	0.613 (0.109)	0.587** (0.016)	0.389* (0.095)	0.670** (0.010)
Interest rate	-0.100 (0.221)	-0.139 (0.139)	-0.082 (0.228)	-0.110 (0.215)	-0.094 (0.350)	-0.207** (0.027)	-0.079 (0.239)	-0.147* (0.077)
Size of welfare state	-0.097 (0.182)	-0.082 (0.298)	-0.082 (0.195)	-0.111 (0.185)	-0.118* (0.065)	-0.166** (0.035)	-0.082 (0.177)	-0.149*** (0.004)
Stock-flow adjustment	0.205* (0.058)	0.147 (0.114)	0.118 (0.151)	0.179 (0.170)	0.212 (0.108)	0.242** (0.010)	0.115 (0.143)	0.216*** (0.006)
Any rule in place	1.734* (0.058)							
Any national fiscal rule		0.866 (0.209)		0.600 (0.399)	1.428 (0.316)			
Stability & Growth pact			1.095 (0.184)	1.145 (0.215)	2.524 (0.126)			
National rule & SGP joint					-2.147 (0.258)			
Strength of national rules						1.431** (0.014)		
Strength of SGP							1.153 (0.161)	
Strength of SGP with national support institutions								1.383*** (0.007)
Observations	108	108	108	108	108	108	108	108
Log-likelihood	-33.71	-36.49	-35.11	-34.68	-33.27	-35.52	-34.90	-32.86
Wald-test of Mundlak averages	0.963	0.967	0.908	0.922	0.862	0.913	0.892	0.958

Note: This table reports the marginal effects of a random effects logit specification, where the marginal effects are calculated at the means of all other variables assuming that the random effects are zero. The Hausman test fails when testing if the random effects specification is valid. Therefore we use the method of Mundlak (1978) and calculate a Wald-test of variables constructed as country specific averages for each explanatory variable in the model. We never reject the hypotheses that these fixed-effects are equal to zero. Therefore the random effects specification is valid. P-values in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

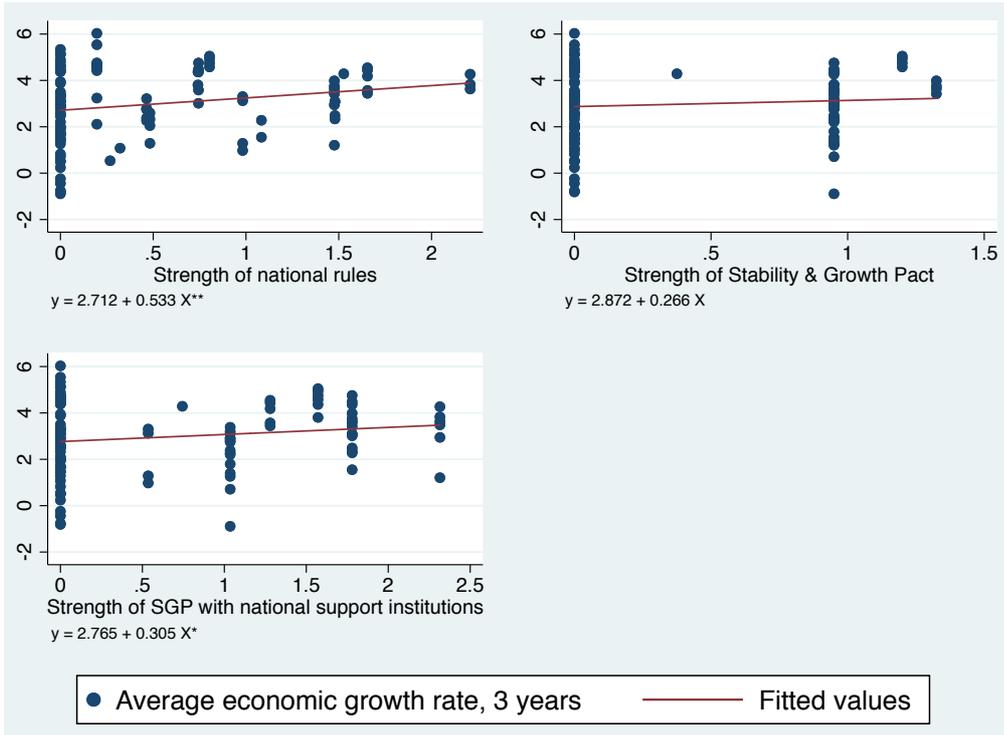
As pointed out by Heineman et al. (2018), government ideology may be correlated with both the debt-to-GDP ratio and fiscal rules. Therefore, we ran the regressions including a measure of government ideology in terms of economic policy manifests. Including this variable does not alter any of our findings. Also, war of attrition within governments may impact both the likelihood of having a rule in place and fiscal outcomes (Alesina and Drazen 1991). However, including a proxy for size fragmentation does not alter our findings. These results are available on request.

Instead of constructing the sample such that only periods with observed preference for fiscal discipline are included, we redo the analysis using all periods where the CAPB is positive. For the 3-year average, this results in a sample of 304 observations and for the 5-year average a sample of 295 observations. For both samples, both the rule dummies and the strength indicators are insignificant using the wild cluster robust standard errors, and only a few are significant at the 10% level using conventional standard errors. If we analyse the full sample of 653 and 611 observations respectively, some the rule dummies even switches sign and are significantly positive for the 3-year average, while neither of the rule dummies are significant for the 5-year averages. For both the 3- and 5-year averages, using the full sample neither of the rule strength indicators are significant. These findings highlight that in some periods in these alternative samples, there exists no clear preference for fiscal discipline, i.e. debt reduction, even while a rule may have been in place. In other words, failing to account for the preference for fiscal discipline distorts the analysis.

5.2 Exploring the credibility mechanism

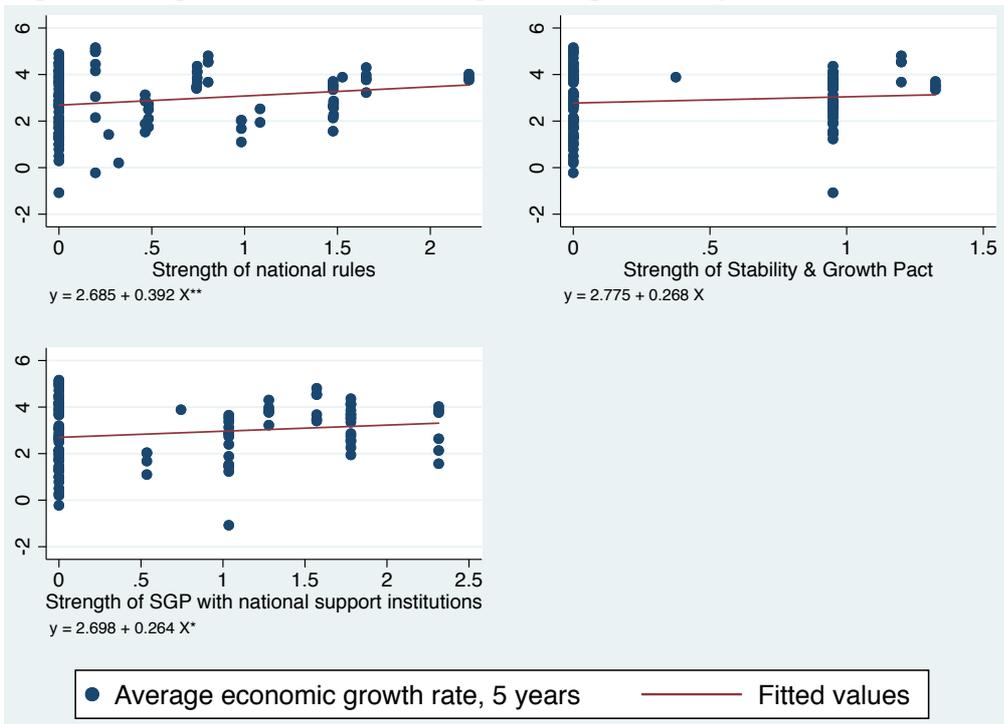
Above, we have argued that the credibility effect of fiscal rules is important in triggering sustained debt reductions. If this is the case we should observe that stronger fiscal rules are associated with higher GDP growth rates. This is indeed the finding that emerges in fig. 3 and fig. 4, which plot average growth rates over 3 and 5 years after a given country-year combination on the y-axis against the three types of fiscal rule strength indexes used in our analysis. Years with a fiscal rule strength value equal to 0 represent years where no rule was in place. Also, as suggested by the analysis above, the credibility effect of the strength SGP measured at the European level is limited. This slope of the bivariate regression line is insignificant concerning the strength of the SGP in both fig. 3 and fig. 4. Whereas national support institutions both concerning national rules and SGP matter for increasing the denominator of the debt/GDP ratio, i.e. it is associated with increased economic activity through a credibility effect as outlined in the theory section.

Fig. 3. Strength of rules and average GDP growth 3 years after



Note: Significance of the slope of the bivariate regression line is indicated in the following way: ** $p < 0.05$, * $p < 0.1$

Fig. 4. Strength of rules and average GDP growth 5 years after



Note: Significance of the slope of the bivariate regression line is indicated in the following way: ** $p < 0.05$, * $p < 0.1$

6. Conclusion

Even though all fiscal rules are associated with debt reduction, some specific measures of fiscal rules cause debt reductions in the longer run, while others do not. Specifically, the mere existence of a national fiscal rule does not robustly lead to sustained public debt reduction during periods with preference for fiscal discipline. On the other hand, the strictness of the national fiscal rules framework, where fiscal rules are both more numerous, more legally binding, have a greater coverage and where compliance with them is being monitored by independent fiscal council, does seem to cause greater levels of sustained debt reductions. In other words, national fiscal rules need to be credible to achieve the goal of debt reduction. Furthermore, supranational fiscal rules, in this case the rules of the European Union's SGP, appear to have caused debt reductions in times with a preference for fiscal discipline. The strength of the SGP itself at the European level does not seem to matter. However, when stronger national fiscal institutions support the SGP, it increases the effect of SGP in terms of causing sustained public debt reductions.

In our view, the results highlight that problems with time inconsistencies play an important role in determining whether rules are effective. Obviously, it is much less economically costly to abandon a national fiscal rule, compared to abandoning the SGP, i.e. leaving the Eurozone. Therefore a country's official status as having signed the SGP may give credibility by itself, in addition to a lasting benchmark effect. When it comes to the strength of the SGP and its effect on debt reduction, our results suggest that the support institutions of the SGP at the European level are largely ineffective, while the presence of strong national institutions supporting it matters. This suggests that the national fiscal institutional anchoring of the SGP matters more compared to the supranational institutional aspects of the rule. Perhaps because this institutional framework is the context in which national governments make fiscal decisions. Overall, the results suggest that the credibility of fiscal rules and institutions are important for their effect on government debt consolidation. Rational economic actors realise the credibility of different fiscal rule arrangements. Consequently, if the rule is credible in the sense that it is less likely to suffer from time inconsistencies both in terms of abandonment and non-compliance, the denominator effect is likely to play a significant role in decreasing the debt-to-GDP ratio.

The results also indicate that a benchmark effect, where supra-national rules serve as a focal point for national governments to reduce excessive deficits may be a cause behind the debt reducing effect. In line with previous findings (Reuter 2015), the effects of fiscal rules might also partly be their ability to function as a benchmark for policy makers and to clarify for them which goals to strive for with regards to fiscal policy. Our results suggest that this benchmark effect might especially kick in during times of fiscal consolidation, or alternatively when voters and/or governments prefer fiscal conservativeness. In these periods, fiscal rules might help policymaker keep an eye on the relevant fiscal aggregates, which help them ultimately achieve a consolidation of public finances. This also seems to suggest that although supra-national fiscal rules such as the SGP might not always be credible enforceable, they might still affect national governments' fiscal policy, as they serve as fiscal benchmarks for national governments and/or are used by these as legitimization for expenditure cuts and tax increases.

However, the heterogenous effects of fiscal rules on government debt found in this article suggest that not all fiscal rules are created equal with regards to their fiscal effects, and that the specific mechanism through which they affect fiscal policy matters. Specifically, the effect of national fiscal rules is, if present, only there in the shorter run. This suggests that these rules suffer from ‘fatigue’ over time, or that national governments might abandon these rules after a short period of time.¹⁹ This is especially the case if these rules are not particular strict and/or embedded in a more stringent institutional framework. These findings lend substantial insights to future design and reform of national fiscal rules and budgetary procedures. National fiscal rules are probably more effective in achieving fiscal consolidation if they are embedded in a stricter legal framework, with greater coverage and are supported by independent auxiliary institutions such as fiscal councils with regards to budget assumptions and rule monitoring.

The surprising large and statistically significant effects of the SGP suggest – in line with recent research (Koehler and König 2015) – that the SGP might actually have had a significant effect on the public finances of the European Union’s member states, despite low official compliance with its numerical public balance and debt limits. The results also suggest that the effectiveness of supranational fiscal rules and institutions might be more effective when there is a stronger national fiscal rules’ framework and vice versa, which is in line with previous recommendation for fiscal institutional reform in the European Union (Hallerberg et al. 2009). It also suggest that the latest approach to fiscal rules in the European Union, exemplified by the 2012 Fiscal Compact, where the treaty specifically states that the supranational fiscal rules of the European Union should be specially implemented in the signatory states’ national fiscal frameworks, might be the right approach to increase the effectiveness of the European Union’s fiscal rules. However, future research and evaluation is needed to make a final verdict of the long-term impact of these changes.

Future research could delve deeper into the role fiscal rules may or may not play. Especially, as previously mentioned, since our empirical strategy has not allowed us to distinguish between potential heterogenous effects of different types of fiscal rules such as expenditure rules, revenue rules, debt rules and balanced budget rules. Thus, it is an interesting research direction to investigate whether these different types of fiscal rules work differently with regards to the success of fiscal consolidations.

¹⁹ Table 2 shows that many national fiscal rules, independent of type, often are in place in shorter intervals. This is clearly opposed to the Stability and Growth Pact, which is in place from 1992 or 1995 onwards for all EU countries.

References

- Aaskoven, L., 2018. "Signaling to Creditors and Voters: The Determinants of the Strengthening of National Fiscal Rules." Working paper, University of Copenhagen.
<https://github.com/Lasseaaskoven/Signalling-to-creditors-and-voters/blob/master/Signaling%20to%20creditors%20and%20voters%2030012018.pdf>
- Alesina, A., Drazen, A., 1991. Why are stabilizations delayed? *American Economic Review* 81, 1170–1188.
- Alesina, A., Perotti, R., 1996. "Fiscal Discipline and the Budget Process." *American Economic Review* 86 (2), 401–07.
- Alesina, A., Ardagna, S., 1998. Tales of fiscal adjustment. *Economic Policy* 27, 489–545.
- Alesina, A., Ardagna, S., 2010. Large changes in fiscal policy: Taxes versus spending. In: Brown, J.R. (Ed.), *Tax Policy and the Economy*, Volume 24, pp. 35–68. Cambridge (MA): National Bureau of Economic Research, Inc.
- Alesina, A., Ardagna, S., 2013. The design of fiscal adjustments. In: Brown, J.R. (Ed.), *Tax Policy and the Economy*, Volume 27, pp. 19–67. Cambridge (MA): National Bureau of Economic Research, Inc.
- Alesina, A., Perotti, R., 1997. Fiscal adjustments in OECD countries: composition and macroeconomic effects. *International Monetary Fund Staff Papers* 44 (2), 210–248.
- Alt, J., Lassen, D. D., Wehner, J., 2014. It Isn't Just about Greece: Domestic Politics, Transparency and Fiscal Gimmickry in Europe. *British Journal of Political Science*, 44, 707-716.
- Andrés, J., Rafael, D., 2006. Automatic stabilizers, fiscal rules and macroeconomic stability. *European Economic Review* 50 (6), 1487-1506.
- Ardagna, S., 2004 Fiscal stabilizations: When do they work and why. *European Economic Review*, 48, 1047-1074
- Asatryana, Z., Castellónb, C., Stratmann, T., 2018. Balanced budget rules and fiscal outcomes: Evidence from historical constitutions. *Journal of Public Economics*, 167, 105-119
- Bai, J., Perron, P., 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* 66, 47–78.
- Bai, J., Perron, P., 2003. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18, 1–22.
- Baldacci, E., Clements, B., Gupta, S., Mulas-Granados, C., 2004. Front-loaded or back-loaded fiscal adjustments: what works in emerging market economies? *IMF Working Paper* 04/157.
- Baldacci, E., Gupta, S., Mulas-Granados, C., 2012. Reassessing the fiscal mix for successful debt

reduction. *Economic Policy*, 365-406.

Bergman, M.U., Hutchison, M., 2015. Economic stabilization in the post-crisis world: Are fiscal rules the answer? *Journal of International Money and Finance* 52, 82-101

Bertola, G., Drazen, A., 1993. Trigger points and budget cuts: Explaining the effects of fiscal austerity. *American Economic Review* 83 (5), 1184–1198.

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust difference-in-difference estimates? *The Quarterly Journal of Economics* 249-275.

Blanchard, O., 1990. Comments on Giavazzi and Pagano. In: Blanchard, O., Fischer, S. (Eds.), *NBER Macro-economics Annual 1990*. MIT Press, Cambridge, MA, pp. 112–117.

Bova, E., Kinda, T., Muthoora, P., Toscani, F., 2015. *Fiscal Rules at a Glance*. Washington D.C.: International Monetary Fund.

Camron, C.A., Gelbach, J.B., Miller D.L., 2008. Bootstrap-Based Improvements for Inference with Clustered Errors, *The Review of Economics and Statistics* 90 (3), 414-427.

Cukierman, A., Tommasi, M., 1998. When does it take a Nixon to go to China? *American Economic Review* 88 (1), 180-97.

Grembi, V., Nannicini, T., Troiano, U., 2016. Do fiscal rules matter? *American Economic Journal: Applied Economics* 8(3): 1–30

Guichard, S., Kennedy, M., Wurzel, E., André, C., 2007. What Promotes Fiscal Consolidation: OECD Country Experiences. *OECD Economics Department Working Papers* 553.

Hallerberg, M., Strauch R.R., von Hagen, J., 2009. *Fiscal Governance in Europe*. New York: Cambridge University Press.

Heinemann, F., Janeba E., Schröder C., Streif F. 2016. Fiscal rules and compliance expectations – Evidence for the German debt brake. *Journal of Public Economics*, 142, 11-23

Heinemann, F., Moessinger, M.-D., Yeter, M., 2018. Do fiscal rules constrain fiscal policy? A meta-regression-analysis, *European Journal of Political Economy*.

Imbens, G. W., Wooldridge, J. M., 2009. Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, 47(1), 5-86.

Kelemen, R.D., Teo, T., 2014. Law, Focal Points, and Fiscal Discipline in the United States and the European Union. *American Political Science Review* 108 (2), 355-370.

Koehler, S., König, T., 2015. Fiscal Governance in the Eurozone: How Effectively Does the Stability and Growth Pact Limit Governmental Debt in the Euro Countries? *Political Science Research and Methods* 3 (2), 329-351.

Krogstrup, S., Wälti S., 2008. Do fiscal rules cause budgetary outcomes? *Public Choice*, 136: 123-138

- Kydland, F.E., Prescott, E.C., 1977. Rules rather than discretion: The inconsistency of optimal plans. *Journal of Political Economy* 85 (3), 473-492
- Lara, A., Wolff, G.B., 2014. Rules and risk in the Euro area. *European Journal of Political Economy*, 43, 222-236
- Larch, M., Turrini, A., 2011. Received wisdom and beyond: lessons from fiscal consolidation in the EU. *National Institute Economic Review* 217 (1), 1-18
- Lavigne, R., 2011. The political and institutional determinants of fiscal adjustment: Entering and exiting fiscal distress. *European Journal of Political Economy* 27, 17–35.
- Lledó, V., Yoon, S., Fang, X., Mbaye, S., Kim, Y., 2017. *Fiscal Rules at a Glance*. Washington D.C.: The International Monetary Fund.
- Mundlak, Y., 1978. On the pooling of time series and cross section data. *Econometrica* 46 (69), 85
- Perotti, R., 1996 Fiscal Consolidation in Europe: Composition Matter. *American Economic Review*, 86 (3), 105-110.
- Poterba, J., 1996. Budget institutions and fiscal policy in the U.S. states. *American Economic Review*, 86, 395–400.
- Reuter, W.H., 2015. National numerical fiscal rules: Not complied with, but still effective? *European Journal of Political Economy* 39, 67-81.
- Robins, J. M., Rotnitzky, A., Zhao, L.P., 1995. Analysis of Semiparametric Regression Models for Repeated Outcomes in the Presence of Missing Data. *Journal of the American Statistical Association* 90(429), 106-121.
- Robins, J. M., Rotnitzky, A., 1995. Semiparametric efficiency in Multivariate Regression Models with Missing Data. *Journal of the American Statistical Association*, 90(429), 122-129.
- Sacchi, A., Salotti, S., 2015. The impact of national fiscal rules on the stabilisation function of fiscal policy. *European Journal of Political Economy* 37, 1–20
- Schaechter, A., Kinda, T., Budina, N., Weber, A., 2012. “Fiscal Rules in Response to the Crisis – Towards the “Next-Generation” Rules. A New Dataset.” *IMF Working Paper 12/187*.
- Tavares, J., 2004. Does right or left matter? Cabinets, credibility and fiscal adjustments. *Journal of Public Economics* 88, 2447-2468.
- van der Laan, M.J., Robins, J.M., 2003. *Unified Methods for Censored Longitudinal Data and Causality*. New York: Springer, Physica-Verlag.
- Wiese R., Jong-A-Pin R., de Haan J., 2018 Can successful fiscal adjustments only be achieved by spending cuts? *European Journal of Political Economy* 54, 145-166.

Wyplosz, C., 2013. Fiscal Rules: Theoretical Issues and Historical Experiences. In *Fiscal Policy after the Financial Crisis*, edited by Alberto Alesina and Francesco Giavazzi, 495-525. Chicago: University of Chicago Press.

Appendix A: Construction of fiscal rules' indexes

National fiscal rules strength index

The construction of the national fiscal rules' strength index, which are inspired by although not completely similar to the methods proposed by Schaechter et al. (2012), is described in detail in Aaskoven (2018) but a basic description of the index construction is given below. First, a strength index is constructed for each type of fiscal rules, expenditure rule, revenue rule, balanced budget rule and debt rule. Each index consists of a simple addition of six to seven indicators for each type of fiscal rule. Each indicator is standardized to run from 0 to 1, where higher values indicate higher levels of rule strength. Not all of the fiscal rules indexes contain all of these indicators and some of the indicators are the same for all types of fiscal rules. The indicators are:

- The legal basis of the rule, which runs from political commitment to coalition agreement to statutory basis and finally to constitutional basis. Having the rule as a purely political commitment gives the score 0.2, having it as a coalition agreement the score of 0.4, statutory status, 0.6, while having the rule as part of the constitutional framework gives the score 1.
- Whether the rule covers the central or general government. Central government gives the score 0.5, while general government gives the score 1.
- The existence of formal enforcement procedures? If yes, then the score is 1, if no, the score is 0.
- The existence of multi-year expenditure ceilings? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules except revenue rule.
- The existence of a Fiscal Responsibility Law? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules.
- Whether an independent body set budget assumptions? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules.
- Whether an independent body monitors budget implementation? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules.

Countries without the fiscal rule in question receive a score of 0 of the strength index. The overall national fiscal rules strength index is then constructed by adding the score for the four fiscal rules strength indexes and standardizing it to run from 0 to 5.

Supranational fiscal rules index (strength of the SGP)

The construction of the supranational fiscal rules index (the SGP strength index) is constructed in a similar way. First, a strength index is first constructed for each type of supranational fiscal rules (expenditure rule, revenue rule, balanced budget rule and debt rule) which consists of an addition of three indicator for each rule. All indicators run from 0 to 1.

- The legal basis of the rule, which runs from political commitment to coalition agreement to statutory basis, to international treaty and finally to constitutional basis. Having the rule as a

purely political commitment give the score 0.2, having it as a coalition agreement the score of 0.4, statutory status, 0.6, having the rule as an international treaty gives the score 0.8, while having the rule as part of the constitutional framework gives the score 1.

- Whether the rule covers the central or general government. Central government gives the score 0.5, while general government gives the score 1.
- The existence of formal enforcement procedures? If yes, then the score is 1, if no, the score is 0.

Then these fiscal rules indexes are added together to form the overall supranational fiscal rules strength index, which is rescaled to run from 0 to 3. It is worth to notice that since the SGP only contains balanced budget and debt rules, all countries in our dataset score a 0 for both the supranational expenditure and revenue rule strength index.

Strength of SGP with national support institutions index

Finally, we construct a joint strength index, where we take into account national auxiliary fiscal institutions, when assessing the strength of the supranational fiscal framework. We do this by again calculating supranational fiscal rules strength indexes for each type of fiscal rule but – like in the case of the national fiscal rules index – we also include indicators for the existence of mainly national auxiliary institutions and rules. Thus, the first three indicators concern the nature of the supranational fiscal rules themselves, while the four other indicators concern auxiliary national fiscal institutions and rules. The indicators are:

- The legal basis of the rule, which runs from political commitment to coalition agreement to statutory basis, to international treaty and finally to constitutional basis. Having the rule as a purely political commitment gives the score 0.2, having it as a coalition agreement the score of 0.4, statutory status, 0.6, having the rule as an international treaty gives the score 0.8, while having the rule as part of the constitutional framework gives the score 1.
- Whether the rule covers the central or general government. Central government gives the score 0.5, while general government gives the score 1.
- The existence of formal enforcement procedures? If yes, then the score is 1, if no, the score is 0.
- The existence of multi-year expenditure ceilings? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules except revenue rule.
- The existence of a Fiscal Responsibility Law? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules.
- Whether an independent body set budget assumptions? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules.
- Whether an independent body monitors budget implementation? If yes, then the score is 1, if no, the score is 0. This indicator is the same for all rules.

After the construction of these four types of joint fiscal rule strength indexes, they are added together form the overall joint strength index, which is rescaled to run from 0 to 6.

Appendix B

Table B1: Explanation of control variables

Variable	Explanation
CAPB	Cyclically Adjusted Primary Balance as share of potential GDP (source OECD)
d.expenditure	Change in government expenditures minus interest payments, as share of GDP
d.revenue	Change in government revenues minus interest payments, as share of GDP
Fiscal impulse	Change in CAPB
Lagged deficit	Lagged CAPB
Lagged debt/GDP	Lagged debt-to-GDP ratio
GDP growth	Real GDP growth rate
Interest rate	Short term interest rate
Size of welfare state	Total government expenditures, minus interest rates, as share of GDP
Stock-flow adjustment	Annual change in gross debt plus the non-cyclically adjusted primary budget balance
Gov. ideology	Continuous from 1-3, one is left-wing, three is right-wing. The variable is weighted average of the ideological composition of the government.
Effective number of gov. parties	Number of parties in government weighted by relative size

Table B2: Descriptive statistics

Variables	Mean	S.D.	Min	Max	Source
Any fiscal rule in place	0.62	0.488	0	1	IMF + self coding
Any national fiscal rule in place	0.463	0.501	0	1	IMF + self coding
Stability and Growth Pact	0.389	0.49	0	1	IMF + self coding
National rule & SGP joint	0.231	0.424	0	1	IMF + self coding
Strength of national rules	0.498	0.634	0	2.208	IMF + self coding
Strength of SGP	0.396	0.499	0	1.325	IMF + self coding
Strength of SGP with national support institutions	0.693	0.801	0	2.315	IMF + self coding
d_expenditure	-0.387	1.19	-3.221	4.326	OECD
d_revenue	0.465	0.959	-2.851	4.504	OECD
Fiscal impulse	1.098	1.583	-5.635	6.971	OECD
Lagged deficit	-0.518	2.771	-9.733	5.358	OECD
Lagged debt/GDP	72.536	29.188	30.226	175.274	OECD
GDP growth	2.786	1.929	-3.374	7.147	OECD
Short term interest	7.049	4.463	0.057	24.9	OECD
Size of government	38.866	8.183	22.976	56.216	OECD
Stock-flow adjustment	0.868	3.668	-18.231	11.434	OECD
Ideology of gov.	2.08	0.819	1	3	DPI + self coding
Effective number of gov. parties	1.705	0.83	1	3.783	DPI + self coding

Note: The statistics are based on the 108 observations that are used in the main analysis. DPI is the abbreviation of Database of Political Institutions from the World Bank

Table B3: Correlation matrix

	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	XII	XIV	XV	XVI	XVII
Any fiscal rule in place	1																
Any national fiscal rule in place	0.73	1															
Stability and Growth Pact	0.62	0.21	1														
National rule & SGP joint	0.43	0.59	0.69	1													
Strength of national rules	0.62	0.7	0.45	0.63	1												
Strength of SGP	0.62	0.25	0.98	0.72	0.48	1											
Strength of SGP with national support institutions	0.68	0.44	0.87	0.74	0.8	0.87	1										
d_expenditure	-0.06	-0.14	-0.08	-0.19	-0.21	-0.1	-0.17	1									
d_revenue	-0.25	-0.22	-0.18	-0.19	-0.25	-0.18	-0.25	0.39	1								
Fiscal impulse	-0.18	-0.14	-0.1	-0.07	-0.12	-0.08	-0.13	-0.23	0.52	1							
Lagged deficit	0.38	0.44	0.28	0.41	0.44	0.29	0.38	-0.05	-0.2	-0.28	1						
Lagged debt/GDP	0.13	0.14	0.23	0.28	0.13	0.18	0.21	0.03	0.05	0.01	0.25	1					
GDP growth	0.14	0.25	-0.03	0.1	0.22	0.02	0.1	-0.63	-0.17	0.05	0.13	-0.11	1				
Short term interest	-0.56	-0.46	-0.34	-0.29	-0.43	-0.33	-0.41	0.29	0.14	0.14	-0.29	-0.14	-0.38	1			
Size of government	0.04	-0.1	0.5	0.42	0.02	0.48	0.33	-0.22	-0.04	0.09	-0.09	0.15	-0.12	-0.12	1		
Stock-flow adjustment	0.07	-0.04	0.13	0.02	-0.14	0.12	0.01	0.09	0.02	0.05	0.16	0.18	-0.28	0.01	0.1	1	
Ideology of gov.	-0.12	-0.2	-0.25	-0.39	-0.36	-0.27	-0.37	0.14	0.12	-0.03	0.04	0.06	0	0.08	-0.27	0.08	1
Effective number of gov. parties	0.05	0.12	0.09	0.18	0.13	0.1	0.1	-0.18	-0.13	0.03	0.14	0.06	-0.1	-0.16	0.2	-0.02	0.05

Note: The correlations are based on the 108 observations that are used in the main analysis

Table B4: Periods with fiscal adjustments, negative debt/GDP growth regimes and successful fiscal adjustments

(1)	(2)	(3)	(4)
Filter specification:	Fiscal adjustment (based on Bai-Perron tests; 5% significance level)	Years with negative growth rate in the debt-to-GDP ratio (Bai-Perron tests 5% significance level)	Periods with successful adjustments based on Bai-Perron approach for both steps
Country and sample length:			
Australia, 1989-2013	1996-98	1996-2008	1996-98
Austria, 1977-2013	1995-97		
Belgium, 1971-2013	1983-87, 1992-98	1994-01, 2002-07	1992-98
Canada, 1970-2013	1986-89, 1995-97	1997-07	1995-97
Denmark, 1973-2013	1984-86, 1998-99	1985-89, 1994-03, 2004-07	1984-86
Finland, 1977-2013	1995-98	1996-2008	1995-98
France, 1978-2013	1995-99, 2001		
Germany, 1970-2013	1981-85		
Iceland, 1980-2013	1990-92, 1995-97	1996-2005	1995-97
Italy, 1970-2013,	1981-83, 1991-93	1979-84, 1999-07	
Japan, 1971-2013	1983-90, 2005-06	1987-1992	1983-90
Netherlands, 1971-2013	1995-97	1994-01, 2002-07	
New Zealand, 1987- 2013		2002-04	
Norway, 1980-2013	1994-97	1979-84, 1994-99, 2007	1994-97
Portugal, 1981-2013	1984, 2009-13	1995-00	
Spain, 1979-2013	1986-87, 1995-99	1999-07	1995-99
Sweden, 1973-2013	1984, 1996-98	1984-90, 1997-13	1984, 1996-98
Switzerland, 1990-2013	1998-99, 2005-08	2005-08	2005-07
United Kingdom, 1972-2013	1979-82, 1995-00	1985-90, 1997-2007	1995-00
United States, 1967-2013	1994-98	1994-00	1994-9
Total no. of years	108		58