# Do Fiscal Rules Constrain Fiscal Policy? A Meta-Regression-Analysis

Friedrich Heinemann, Marc-Daniel Moessinger, and Mustafa Yeter

# ZEW

Zentrum für Europäische
Wirtschaftsforschung GmbH

Centre for European
Economic Research

Discussion Paper No. 16-027

# Do Fiscal Rules Constrain Fiscal Policy? A Meta-Regression-Analysis

Friedrich Heinemann, Marc-Daniel Moessinger, and Mustafa Yeter

# Do Fiscal Rules Constrain Fiscal Policy?
# A Meta-Regression-Analysis

Friedrich Heinemann[1]

Marc-Daniel Moessinger[2]

Mustafa Yeter[3]

March 21, 2016

## Abstract

We implement a meta-regression-analysis for the budgetary impact of numerical fiscal rules. Based on 30 studies published in the last decade, we offer a consensus estimate with respect to the level of statistical significance, provide suggestive evidence for the effect size, and identify study features of relevance for the measured impact of fiscal rules. Overall, the results document a constraining impact of rules. However, this impact is weakened if refined identification strategies are employed. Moreover, the results provide evidence for a publication bias in which journals are more likely to report constraining and statistically significant effects compared to working papers. We further provide recommendations for future research on the budgetary impact of fiscal rules.

**JEL classification:** H50; H6; H63

**Keywords:** fiscal rules; meta-regression-analysis; public finances

[1]ZEW Mannheim and University of Heidelberg, E-mail: `heinemann@zew.de`

[2]ZEW Mannheim, E-mail: `moessinger@zew.de`

[3]Corresponding author; ZEW Mannheim, P.O. Box 103443, 68034 Mannheim, Germany. E-mail: `yeter@zew.de`

# 1    Introduction

In recent decades, industrialised countries have exhibited a substantial increase in public debt reaching levels which increasingly threaten fiscal sustainability. There is a growing consensus that this is due to a deficit or spending bias of politicians which distorts democratic budgetary decision making (see Imbeau, 2004, for a survey). Legislators in numerous industrialised countries therefore proposed and introduced numerical fiscal rules in order to meet these adverse incentives as well as the long-run trend in increasing debt. Under the impression of the recent crisis, the introduction and strengthening of these rules ranked high on the political agenda. At the supranational level, for instance, the European Union has strengthened the Stability and Growth Pact with stricter thresholds and more stringent decision making procedures. Furthermore, the Fiscal Compact obliged ratifying EU Member States to introduce numerical fiscal rules or to adapt existing ones along the requirements of the compact in order to increase their efficacy. The effectiveness of fiscal rules in limiting public debt, however, has been a matter of discussion.

In recent years, a large body of empirical research investigating the effectiveness of fiscal rules has emerged.[1] However, a majority of these studies share a common point of criticism as they neglect or do not deal adequately with the potential issue of endogeneity. The latter might invalidate the empirical analysis, as the fact that a country has a fiscal rule in place might primarily reflect its preferences for fiscal discipline (Poterba, 1996; Debrun et al., 2008). The establishment of a fiscal rule or its strictness might systematically coincide with the fiscal condition of a particular country. Countries already exhibiting a low level of public debt might therefore be more likely to coincide with stricter rules. Consequently, a common cause interdependence may exist in which preferences drive both the presence of fiscal rules as well as the fiscal performance of a country. Thus, in these kinds of settings, fiscal rules cannot be interpreted as exhibiting a causal influence on the level of public debt.

Against this background, our contribution is twofold. First, we conduct the first meta-regression-analysis (MRA) on this particular field of literature summarising existing empirical evidence in a systematic way. To this end, we concentrate on the impact of *numerical fiscal rules* on fiscal aggregates in a national as well as subnational context.[2] Second, we explicitly account for different identification strategies followed by primary studies and assess the relevance of the above mentioned endogeneity concerns.

Given the large methodological heterogeneity in this field of literature, several challenges have to be overcome in order to implement a MRA. As fiscal rules are measured in various ways, the extraction of a uniformly defined and comparable impact coefficient is

---

[1]    Early studies usually focused on balanced budget rules in the context of US states (e.g., Bails, 1982; Abrams and Dougan, 1986; Shadbegian, 1996). Subsequently, analyses have been extended to further countries or even further lower tier subnational jurisdictions (for qualitative surveys see Kirchgässner, 2002; Burret and Feld, 2014).

[2]    As a consequence, we preclude studies focussing exclusively on the impact of supranational rules or budgetary procedures (e.g., von Hagen, 2002)

not feasible. To this end, we follow meta-analytical studies such as Card et al. (2010) and rely on the *t*-value rather than the point estimate itself. This measure is fully comparable across studies and indicates the sign as well as the level of statistical significance with which fiscal rules impact on fiscal aggregates. For the largest homogeneous sub-group of studies, we additionally run meta-regressions with respect to the effect size. Furthermore, given the wide range of control variables employed in primary studies, we subsume particular features in groups of control variables which define a typical study and against which we test the robustness to alterations in the estimation specifications in existing studies.

Based on 30 studies which have been published in the period between 2004–2014, the following insights emerge: In general, the MRA provides evidence for a constraining and statistically significant impact of fiscal rules on fiscal aggregates at the national level. This result is particularly strong for deficits and to a much lesser degree for debt, expenditures or revenues. Furthermore, the results point to a substantial bias if the potential endogeneity of fiscal rules is neglected, confirming that these concerns have to be taken seriously. The statistical significance of the impact of rules is significantly reduced if a primary study is characterised by a more sophisticated identification strategy. With respect to the administrative level, we find evidence for a higher efficacy of fiscal rules at the municipal level compared to the national level. Beyond that, the MRA provides evidence for a potential publication bias as results obtained from working papers are on average associated with lower levels of statistical significance. With respect to the effect size, our MRA points to a deficit reducing impact of fiscal rules in the range of -1.5 to -1.2% of GDP.

In the next section, we describe our algorithm for the literature search. In section 3, we explain the coding of the meta-variables and take a descriptive look at our sample. While the empirical methodology is described in section 4, the meta-regression results are presented in section 5. Section 6 concludes.

## 2   Selection of Studies

Crucial for the validity of a MRA is the application of a well-defined and reproducible algorithm guiding the literature search as well as the coding for the subsequent econometric modelling. To this end, we base our analysis on the guidelines proposed by Stanley et al. (2013).[3]

We conducted a literature search using the *EconLit* database. This database covers all economic journals indexed by the American Economic Association and provides a comprehensive overview of the economic literature. The keywords employed to conduct the literature search involved *fiscal rule*, *fiscal restraint*, *debt brake*, *debt rule*, *budget rule*, and *deficit rule*. The search was carried out on April 16th, 2014 and repeated on January

---

[3]     Further guidelines for meta-analyses are offered by Stanley and Doucouliagos (2012) or Nelson and Kenedy (2009).

28$^{\text{th}}$, 2015 to ensure that all publications from 2014 were included. In raw figures, this resulted in a total of more than 1,500 search results.

We then filtered out all duplicate search results found across keywords and excluded papers not published in English. Furthermore, we limited the time frame of the search to 10 years, i.e., only studies published between 2004 and 2014 were included. Consequently, our MRA provides a consensus estimate formed by the empirical evidence from the last decade.[4]

To derive a sample of comparable studies, we further applied a common definition of a *fiscal rule*. To this end, we rely on the definition provided by Kopits and Symansky (1998) where fiscal rules are characterised as a permanent numerical constraint on fiscal policy defined in terms of an indicator of overall fiscal performance such as the government deficit, debt or expenditure. Against this background, we focus in our MRA on *numerical fiscal rules* in national and subnational contexts, and preclude studies which exclusively investigate the impact of supranational rules such as the Stability and Growth Pact. This decision is guided by the usual focus on cross-jurisdictional variation in rules in the empirical literature.[5]

Finally, we limit our analysis to primary studies with public finance indicators such as expenditures, revenues, debt, or deficit as dependent variables including their overall state or their subcomponents. In contrast, we exclude studies testing the impact of fiscal rules on financial market indicators or the structure of the budget. The latter do not fit into the framework of our research question as they lack a clear hypothesis against which the result with respect to the efficacy of a fiscal rule could be interpreted (e.g., Dahan and Strawczynski, 2013).

Altogether, this algorithm yields 51 studies for coding. In assessing and coding these studies, we follow a four-eyes principle to reduce measurement error and imprecision caused by interpretation. Initially, each paper is read by one of the authors. This involves a first assessment in terms of relevance with respect to the guidelines defined above. If these are fulfilled, the same author codes all relevant features of the primary study with respect to the dimensions included in the MRA. In a subsequent step, each paper is once again read individually by one of the remaining authors and coding decisions are double checked. This results in a total of 30 studies forming the basis of our analysis. Table A.1 provides a comprehensive overview of these papers.

---

[4]    Please note that this does not apply to the time coverage of primary studies.

[5]    While we preclude studies focusing exclusively on supranational rules, the latter is taken explicitly into account in our MRA if they have been used as a control variable (e.g., Dahan and Strawczynski, 2013).

# 3 Data and descriptive statistics

## Variable coding

In order to render the estimation of a meta-regression feasible, idiosyncratic features of the considered primary studies must be merged into a set of common and comparable control variables. This process comes along with several key challenges. In the context of a MRA, one would ideally like to extract a uniformly defined impact coefficient which would be immediately comparable across primary studies. In our context, however, this is not feasible due to the large methodological heterogeneity within this strand of literature.

First, there is substantial heterogeneity with respect to the dependent variable in primary studies. While slightly less than one third of our observations relate to the primary deficit as dependent variable, one quarter refers to either expenditures or revenues. Also the coding of the dependent variable differs, i.e., dependent variables are expressed in relation to the GDP, revenues, or in per capita or absolute terms. Together these circumstances impede the direct comparability of effects across studies.

Second, there is substantial heterogeneity with respect to the codification of fiscal rules. As shown in section 3, primary studies in our MRA capture the impact of fiscal rules either through dummy, discrete, or continuous variables. Given these differences, comparing marginal effects of fiscal rules across studies is not possible as the switch from 0 to 1 in case of a dummy variable is not necessarily comparable to a change from an index value $a$ to $b$ in case of a continuous index. Therefore, we follow meta-analytical studies such as Card et al. (2010) which face a similar challenge and which rely on the $t$-value rather than on the coefficient of the impact of fiscal rules.[6] The $t$-value is immediately and fully comparable across studies and indicates whether a fiscal rule has a positive, negative, or no statistically significant impact on the respective fiscal indicator.

A clarification is necessary with respect to the sign convention applied in the context of our MRA. In light of the core research question, the goal of this MRA is to subsume existing empirical evidence on whether fiscal rules *constrain* budgetary policy and shift it towards a more sustainable path. If fiscal rules are indeed effective in this regard, they should *reduce* deficits, debt as well as expenditures and *increase* revenues. In order to account for these aspects, we code $t$-values of primary studies in a way that a negative sign always indicates a constraining impact, i.e., a fiscal rule *lowers* deficit, debt, and expenditures, or *increases* revenues, respectively.[7]

We account for the richness of specifications across considered primary studies by subsuming specific dimensions into the following variable groups: (i) the *classification*

---

[6]    Further studies relying on $t$-values instead of coefficients involve Baskaran et al. (2014) as well as Klomp and de Haan (2010).

[7]    In case of primary studies relying on revenues as the dependent variable, we invert the sign of the $t$-values capturing the impact of fiscal rules. By the same token, we also invert the sign of studies using the budgetary balance as the dependent variable. Yet, this does not apply to studies using the deficit.

*of the dependent variable*, (ii) the *type of the fiscal rule*, (iii) the *administrative level* the primary study's data is referring to, (iv) a general set of *control variables* usually employed in studies within this strand of literature, (v) indicators testing for a potential *publication bias*, and (vi) characteristics of the *econometric specification*. Additionally, we pay particular attention to strategies used to deal with the issue of endogeneity and subsume respective strategies in the group (vii) *identification strategies*. Within these groups, usually several individual variables account for different aspects of the respective topic. In order to test for the robustness of our results, we extend the set of groups and include the *country* as well as *time coverage* of primary studies. A detailed description of all meta-analytical variables is provided in Table A.2 in the appendix.

The group of *identification strategies* needs a more detailed discussion. While there exists a large body of empirical research investigating the effectiveness of fiscal rules, a majority of these studies share a common point of criticism. Many primary studies neglect or only deal insufficiently with the potential presence of a common cause interdependence, omitted variables bias, or reversed causality. Some empirical studies tried to tackle these potential issues. For instance, one strand of the literature argues in favour of a time-invariance of fiscal preferences. Given such a setting, implementing a simple fixed effects estimation accounting for unobserved and time-invariant individual heterogeneity would suffice to take the impact of fiscal preferences into account (e.g., Krogstrup and Wälti, 2008). Another strand of the literature relaxes the assumption of time-invariant preferences and tries to measure fiscal preferences explicitly.[8] Primary studies implementing IV-strategies or quasi-experimental methods pose another approach with respect to the causal inference. For instance, Foremny (2014) implements an IV-estimation relying on political determinants of the level of government as well as indicators for the general fiscal stance of a country as instruments. Grembi et al. (2012) applies a difference-in-discontinuities approach in a subnational context for Italian municipalities.

In order to account for these differences in identification approaches, we derive four different dummy variables which subsume different dimensions of the described strategies. The *Identification dummy 1* forms the most comprehensive approach and is set to one if at least one of the above described four strategies (fixed effects, measurement of fiscal preferences, IV-strategies, and quasi-experimental methods) are employed in the primary study. The subsequent dummies each drop a particular strategy sequentially. While the *Identification dummy 2* excludes jurisdiction fixed effects from its definition, *Identification dummy 3* subsumes IV-strategies as well as quasi-experimental methods, whereas *Identification dummy 4* is only set to 1 if quasi-experimental methods are implemented.

---

[8]    Heinemann et al. (2014), for instance, use historical inflation rates or some measure of political preferences of the executive government.

## Descriptive statistics

In the following, we present specific sample characteristics. In order to avoid that the results suffer from extreme outliers, we trim the distribution of $t$-values by excluding the top and bottom 1% percentiles.[9] In all summary tables, we present the number of observations and the percentage share in relation to the total number of 1,033 observations, the number of studies employing the respective variable, minimum and maximum $t$-values as well as unweighted and weighted mean $t$-values. The latter are computed as the number of observations extracted from primary studies are not equally distributed. For instance, just two studies out of 30 (Nerlich and Reuter, 2013: 215 observations; Tapsoba, 2012: 220 observations) account for almost half of the full sample (435 observations in relation to 1,033 observations). In order to avoid biased descriptives or point estimates (section 5), we apply an analytical weighting procedure and weight each observation with the inverse of the share of observations per study in relation to the full sample. As a result, we assign an equal weight to each study in the full sample.[10]

Table 1: Summary for the classification of the dependent variable

| Variable | Obs. | % sample | Studies | Mean $t$-value | | Min. | Max. |
| --- | --- | --- | --- | --- | --- | --- | --- |
| | | | | unweighted | weighted | | |
| Debt | 119 | 11.5 | 7 | −1.538 | −1.273 | −4.710 | 3.360 |
| Secondary deficit | 128 | 12.4 | 15 | −2.116 | −2.134 | −10.630 | 1.900 |
| Primary deficit | 280 | 27.1 | 13 | −2.265 | −1.778 | −9.890 | 2.498 |
| Secondary exp. | 221 | 21.3 | 13 | −1.647 | −0.874 | −11.730 | 4.311 |
| Primary exp. | 125 | 12.1 | 4 | −1.931 | −1.720 | −4.042 | 4.038 |
| Revenue | 160 | 15.4 | 10 | 0.024 | −0.155 | −9.400 | 4.250 |

Notes: The total number of observations is 1,033 based on 30 studies. The statistics are based on trimmed data excluding the top and bottom 1% percentiles. Observations are weighted with the inverse of the share of observations per study in relation to the full sample. Because several studies employ more than one dependent variable, the sum of studies exceeds the total number of studies.

Table 1 gives a summary of the *classification of the dependent variable.* We distinguish between six dependent variables: debt, secondary deficit, primary deficit, secondary expenditure, primary expenditure, and revenue. Most of the studies rely on the deficit (39.5% of the total) or expenditures (33.5% of the total) as the dependent variable. On average, observations taken from studies with debt as the dependent variable lack statistical significance at conventional levels. The opposite holds true for observations obtained from studies relying on the deficit. Irrespective of the weighting, observations in studies with secondary and primary deficit as the dependent variable are on average statistically significant at the 5% level. In contrast, the weighting reduces the statistical precision of studies with secondary expenditures as a dependent variable. For all dependent vari-

---

[9]    Without trimming, the minimum and maximum $t$-values are equal to -40.710 and 11.230, respectively. The $t$-value distribution distinguished along types of dependent variables and without trimming as well as weighting is shown in Figure A.1 in the appendix.

[10]   For more information see section 4

ables with the exception of revenues, mean $t$-values exhibit a negative sign indicating a constraining effect of rules on average irrespective of weighting.

To gain further insights into the distribution of $t$-values, we also present a boxplot graph in Figure 1. The dashed line marks the 5% threshold for statistical significance ($t$-value = –1.96). The margins of the boxes indicate the 25% and the 75% quartiles whereas the vertical line in between depicts the median $t$-value.[11] In particular studies employing the secondary deficit as dependent variable exhibit statistically significant results. More than 50% of the observations are statistically significant at the 5% level. A similar result is found for studies relying on the primary deficit or primary expenditure with a median $t$-value being only slightly below the 5% threshold. In case of debt as the dependent variable, only 25% of the observations are statistically significant at the 5% level. Concerning the skewness of the distribution, $t$-values from observations with debt and (to a lesser extent) primary deficit as dependent variable are skewed to the left while the opposite holds true for observations with primary and secondary expenditure.

Figure 1: Boxplot for the classification of the dependent variable
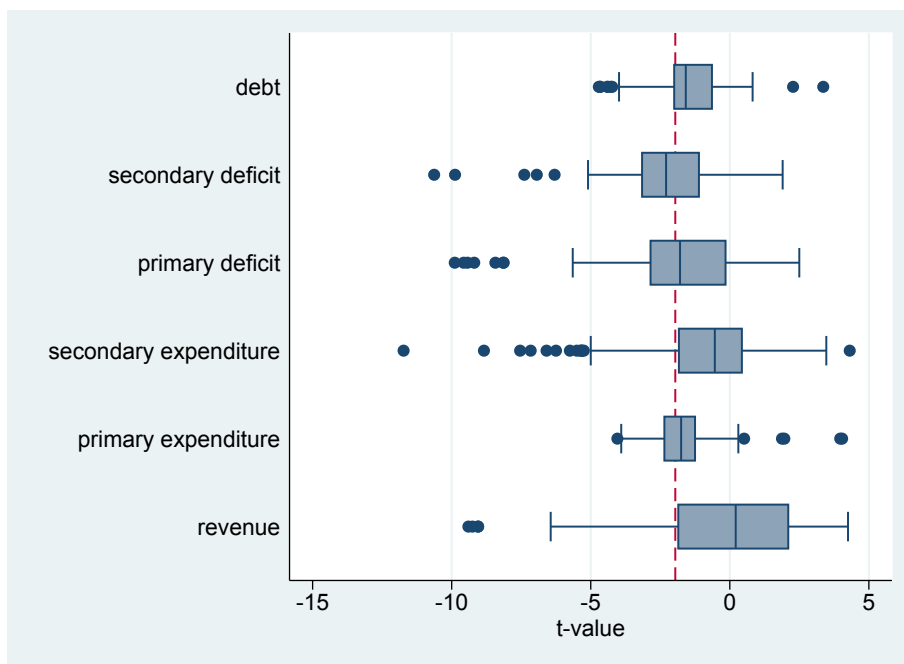


Notes: The total number of observations is 1,033 based on 30 studies. The statistics are based on trimmed data excluding the top and bottom 1% percentiles. Observations are weighted with the inverse of the share of observations per study in relation to the full sample. The dashed line marks the 5%-threshold for statistical significance ($t$-value = –1.96).

In Table 2, we present summary statistics for the *type of the fiscal rule* and the *administrative level* to which the primary studies' data refer to. With respect to the former, we distinguish between debt rules, deficit rules, expenditure rules, and revenue

---

[11]    Single dots represent outliers identified by standard whisker plots using a distance of 1.5 times the inter-quartile range above the 75% quartile and below the 25% quartile.

rules.[12] Again, there are differences between weighted and unweighted means. While most of the unweighted mean $t$-values of various rule types indicate a statistical significance at the 10% level (the exception is the debt rule which is marginally insignificant), statistical significance is lost when we employ analytical weights. However, observations in case of deficit rules are close to statistical significance with a weighted mean $t$-value of -1.515. With respect to the distribution of rule types across studies, deficit rules are used in more than 70% of the cases, while revenue rules are only employed in about one third of the sample. Mean $t$-values and information on the distribution of observations for the *administrative level* are presented in the bottom part of Table 2. In total, 63.2% of the observations stem from studies for the national level, while the remaining part is equally shared between observations from the state or municipal level. The most stable (with respect to weighting) and statistically significant effect is found for observations with data for the municipal level.

Table 2: Summary for the type of fiscal rules and the administrative level

| Variable | Obs. | % sample | Studies | Mean $t$-value | | Min. | Max. |
|---|---|---|---|---|---|---|---|
| | | | | unweighted | weighted | | |
| *Type of fiscal rule* | | | | | | | |
| Debt rule | 485 | 46.9 | 16 | $-1.629$ | $-1.033$ | $-9.890$ | 4.311 |
| Deficit rule | 753 | 72.8 | 27 | $-1.732$ | $-1.515$ | $-11.730$ | 4.311 |
| Expenditure rule | 496 | 45.4 | 20 | $-1.697$ | $-0.932$ | $-9.890$ | 4.311 |
| Revenue rule | 339 | 32.8 | 15 | $-1.723$ | $-0.973$ | $-9.890$ | 4.311 |
| *Administrative level* | | | | | | | |
| National | 653 | 63.2 | 17 | $-1.649$ | $-1.124$ | $-9.890$ | 4.311 |
| State | 192 | 18.5 | 8 | $-1.400$ | $-1.297$ | $-5.097$ | 4.250 |
| Municipal | 188 | 18.1 | 5 | $-1.829$ | $-2.154$ | $-11.730$ | 3.470 |

Notes: The total number of observations is 1,033 based on 30 studies. The statistics are based on trimmed data excluding the top and bottom 1% percentiles. Observations are weighted with the inverse of the share of observations per study in relation to the full sample. Because several studies employ more than one type of a fiscal rule, the sum of studies and observations exceeds the total number of studies/observations.

We present information on the coverage of control variables and applied identification strategies in Table 3. In contrast to the tables presented above, only weighted mean $t$-values distinguishing between studies including this particular variable (or identification strategy) or not are shown. Furthermore, we present the number and percentage shares of observations and studies. The inclusion of specific groups of controls has an effect on the statistical significance of fiscal rules. The mean $t$-value becomes statistically significant if authors control for ideology, government stability, electoral cycles, or supranational rules such as the Stability and Growth Pact. In contrast, the results are only marginally affected by controlling for debt, openness, or the dependency ratio.

With respect to the application of identification strategies, we distinguish between the inclusion of jurisdiction fixed effects, proxies measuring preferences and cultural as-

---

[12] As some primary studies employ composite indicators which comprise several rule types, the sum of the percentage shares exceeds 100%.

Table 3: Coverage of control variable groups and identification strategies

| Variable | Obs. | % sample | Studies | % studies | Mean $t$-value if | |
|---|---|---|---|---|---|---|
| | | | | | included | not included |
| *Groups of control variables* | | | | | | |
| Ideology | 441 | 43 | 12 | 40 | -1.958 | -1.085 |
| Government stability | 316 | 31 | 4 | 13 | -2.214 | -1.228 |
| Election year | 526 | 51 | 11 | 37 | -1.677 | -1.161 |
| Debt level | 610 | 59 | 13 | 43 | -1.189 | -1.444 |
| Openness | 265 | 26 | 6 | 20 | -0.444 | -1.527 |
| Dependency ratio | 481 | 47 | 4 | 13 | -1.488 | -1.321 |
| Supranational rule | 557 | 54 | 6 | 20 | -2.009 | -1.195 |
| *Identification strategy* | | | | | | |
| Jurisdiction fixed effects | 690 | 67 | 25 | 83 | -1.459 | -0.939 |
| Preferences and culture | 171 | 17 | 8 | 27 | -1.778 | -1.256 |
| 2SLS IV | 28 | 3 | 5 | 17 | -0.544 | -1.387 |
| Quasi-experimental | 317 | 31 | 4 | 13 | -1.563 | -1.317 |

Notes: The total number of observations is 1,033 based on 30 studies. Only weighted observations are presented. The statistics are based on trimmed data excluding the top and bottom 1% percentiles.

pects directly, and the application of IV-estimators or quasi-experimental designs. For the majority of identification strategies, their application is associated with statistically insignificant average $t$-values for the impact of fiscal rules. The sole exception is given by proxies for preferences and culture. When authors account for these preferences, average $t$-values range at the 10% level of statistical significance. However, larger absolute $t$-values, i.e., more constraining effects of fiscal rules are observed if authors apply identification strategies. Here, the sole exception is given by the application of IV-estimators.

# 4   Empirical Methodology

In contrast to classical meta-analyses, we rely on MRA techniques. While the former combines individual study results to an overall or consensus estimate by assuming between-study homogeneity, the latter explicitly investigates the extent to which heterogeneity in primary results can be explained by one or several characteristics of the respective studies (Harbord and Higgins, 2008).

As explained above, due to the heterogeneity in coefficients across primary studies, we employ the $t$-value of the fiscal rule's impact as our meta-analytical dependent variable. In line with the usual assumptions in case of MRA, we expect that heterogeneity in individual results of primary studies is driven by sampling error ($\epsilon_{s,i}$), study- or design-specific factors ($\mathbf{X}$) and the set of estimate-specific covariates ($\mathbf{Z}$). Our baseline estimation can therefore be summarised by equation (1).

$$\hat{t}_{s,i} = t_0 + \mathbf{X}_{s,i}\beta + \mathbf{Z}_{s,i}\gamma + \epsilon_{s,i} \tag{1}$$

Here, $\hat{t}_{s,i}$ refers to the estimated $t$-value in case of specification $i$ from primary study $s$. $t_0$ is the intercept and captures the respectively defined baseline $t$-value of the typical study across all primary studies considered in our MRA. This baseline is defined as an estimation of the impact of a *deficit rule* on the *primary deficit* for the *national level*. Furthermore, the statistical significance of the respective coefficient is computed with *robust standard errors* and the study is published in a *refereed journal.* The definition of a baseline is necessary due to the way of coding of the study- or design-specific features in the context of a MRA. This is done by individual dummy variables which are then usually combined to a sub-group in order to capture a particular design feature of the primary study. In this process, dummies constituting such a sub-group are usually self-excluding. Consequently, considering all of these dummy variables at the same time in our meta-regression would result into perfect multi-collinearity. Therefore, we define one particular feature as baseline for every sub-group where self-exclusion applies. Taken together, all these baseline features are captured by the intercept of our MRA. All reported coefficients have to be interpreted relative to this baseline and report an impact of the deviation in this particular feature from the baseline.

Given the heterogeneity in methodological approaches and sample sizes in primary studies, the variances of individual estimates of $\hat{t}_{s,i}$ are not homoscedastic. Consequently, the estimation of equation (1) will be more efficient using *weighted least squares* instead of ordinary least squares (Greene, 2003). The selection of proper weights, however, is not straightforward as our MRA does not rely on standardised coefficients. While the literature usually uses the inverse of error term variances obtained from primary studies, this is not feasible in case of a MRA using $t$-values. Since we follow a multiple sampling approach in which information from all available estimations in primary studies are used, the number of observations resulting from particular studies differ substantially. In order to avoid settings in which results and design features of particular primary studies constituting a large share in our sample drive the results of our MRA, we use analytical weights based on the inverse of the share of observations per study in relation to the full sample. These weights assign an equal weight to every primary study in our estimation sample.

The fact that usually more than just one observation is obtained from each individual study has a further impact on the estimation of equation (1). As these individual observations are likely to be correlated due to the low within-variation in design features, error terms in the estimation of equation (1) are likely to be correlated as well. We therefore cluster standard errors at the level of studies considered in our MRA. However, since our sample consists of 30 studies, clustering will result in further econometric issues. First, one may argue that the number of clusters is too low potentially resulting in an under-estimation of the underlying serial correlation in individual observations (Angrist and Pischke, 2009). Second, given the large heterogeneity in the number of primary estimates obtained from particular studies, the size of clusters will vary accordingly. Therefore, we also apply the so called *wild-cluster-bootstrap procedure* in order to deal with both

issues (Cameron et al., 2008). This permits to implement a bootstrapping procedure for the standard errors of equation (1) while taking the clustered structure of the data into account and re-sampling along them (Cameron and Miller, 2015).

We extend and vary the estimation methodology of equation (1) in several ways in order to test for the robustness of our results. First, we include two more groups of control variables in our MRA, namely the *country* and *time coverage* of primary studies. While this extension provides further insights compared to the baseline specification, it comes at the expense of dropping the administrative level due to the highly pronounced correlation with these additional dimensions. In general, we also report the variance-inflation-factor in order to test for a potential bias due to high correlation between control variables.

A further extension relates to the analysis of the *publication bias*. In a first step, we change the baseline of this group and test for the impact of journal publications explicitly. In a second step, we use information about the impact factor of journals included in our MRA. The latter may allow to deduce a differential publication bias which may be dependent on the reputation of the particular journal the primary study has been published in.[13]

With respect to the estimation methodology, we implement two different approaches. First, we extend equation (1) by adding study fixed effects. This allows to control explicitly for unobserved study specific aspects which are by definition *time-constant* in the context of a MRA. However, implementing a fixed effects regression increases the demands on the control variables in a MRA as they must exhibit within variation. More precisely, this means that dimensions we control for in our MRA must vary across specifications within a particular study. As this is not necessarily the case for all dimensions in all primary studies, the meta-regressions employing fixed effects can only be implemented on a reduced set of control variables.

Second, we follow Card et al. (2010) and implement an ordered probit approach as a further extension to the estimation methodology. To do so, $t$-statistics obtained from primary studies are split into three groups according to equation (2). These groups involve *negative significant* (-1), *insignificant* (0), and *positive significant* (1) observations. The distinction is made by using the critical value of the $t$-distribution at the 5 % threshold. Obtained results are interpreted directly with respect to their sign and level of statistical significance instead of computing explicit marginal effects (see Card et al., 2010; Cameron and Trivedi, 2010).

$$\tilde{t}_{s,i} = \begin{cases} -1, \text{ negative significant} & \text{if} \quad \hat{t}_{s,i} \leq -1.96 \\ \phantom{-}0, \text{ insignificant} & \text{if} \quad -1.96 < \hat{t}_{s,i} < 1.96 \\ \phantom{-}1, \text{ positive significant} & \text{if} \quad \hat{t}_{s,i} \geq 1.96 \end{cases} \tag{2}$$

---

[13] While we have explicit panel information about the impact factors of journals considered in our MRA, we do not have respective information for working papers. In case of the latter, we assign a zero impact factor to working papers considered in our MRA.

Finally, we also run estimations for the effect size. For this analysis, the sample has to be limited to the largest homogenous subgroup which comprises of studies testing for the impact of fiscal rules on both primary and secondary deficit expressed relative to the GDP as well as rules measured in terms of dummy variables. As in case of the baseline, we weight observations and estimate using weighted least squares. However, given the reduced sample size, the estimation specification must be altered. This is due to the lack in coverage of particular dimensions as well as resulting collinearity between remaining groups of control variables in the MRA.[14] Given the reduced sample size as well as reduced number of considered primary studies, results from this part of our analysis should only be interpreted with caution and seen as suggestive.

# 5   Results

## Main results

Table 4 summarises our baseline estimation results. All five columns report estimations of equation (1) using weighted least squares. The sample constitutes of 30 primary studies from which 1,033 observations are obtained. We report both clustered as well as wild-cluster-bootstrapped *t*-values. While the former is noted in brackets, braces are used in case of the latter.

Our MRA puts a special focus on the issue of endogeneity within this field of literature and the identification strategies followed in primary studies. Results in column (1) serve as the main specification in which components constituting the identification dummies are used individually. Columns (2) to (5) report results using sequentially the set of four identification dummies derived in section 3.

Table 4 is dived into subsections by headlines indicating the respective group of study and design dimensions as well as their respective baseline category. The definition of the baseline study is constant throughout all specifications. It uses the *primary deficit* as dependent variable, captures the impact of a *deficit rule*, relies on data for the *national level*, reports *robust standard errors*, and is published in a *refereed journal*.

In case of the main specification, the average consensus *t*-value for this baseline study is equal to -1.874. More precisely, primary studies exhibiting the design features defined above are found to report a negative and statistically significant impact of fiscal rules at the 10% threshold on average. Except for columns (2) and (5), this holds true for all other specifications in Table 4. The baseline in case of column (5) even indicates a negative and statistically significant impact at the 5% level. These effects are estimated with statistical precision in our MRA in case of both clustered as well as wild-cluster-

---

[14]    We explicitly control whether all remaining variables in the MRA for the effect size are backed by more than one study. This is especially important given the reduced sample size and reduced number of considered studies.

bootstrapped standard errors.

Relative to the baseline, variation in the *dependent variable* matters for the impact of fiscal rules. This is the case for debt, primary as well as secondary expenditures, and revenues. All variables are associated with positive coefficients in case of our MRA, hereby indicating a reduction in the level of statistical significance with which fiscal rules affect these fiscal aggregates. For instance, in case of debt, the results of the MRA indicate a range of average *t*-values from -0.729 to 0.113. As a consequence, these results imply a level of significance below the 10% threshold. This general observation holds true for all *dependent variables* for which coefficients are estimated with statistical precision in our MRA. With the exception of primary expenditures, coefficients are estimated with statistical precision at the 1% level in case of both clustered as well as wild-cluster-bootstrapped standard errors.

In contrast to the *dependent variable*, variation in the *type of fiscal rules* does not matter and does not seem to exhibit an additional impact. This is the case throughout all five specifications and may be driven by the fact that numerous primary studies do not capture the type of the rule precisely or lump it together with other types of rules.

New insights can be gained from the analysis of the *administrative level* the data of primary studies refer to. Here, the results of our MRA indicate that on average and in absolute terms higher *t*-values are associated with studies relying on data referring to the municipal level. This indicates that relative to the baseline which captures a significant impact at the national level, rules at the municipal level exhibit an even stronger statistically significant impact. The results imply levels of significance at the 1% threshold and are estimated with statistical precision in our MRA.

With respect to *control variables*, the picture is less clear cut. Usual control variables employed in primary studies such as the debt level, government stability or election years are not associated with an additional impact in terms of statistical significance. Other variables such as the ideology or dependency ratio exhibit a differential impact which, however, is not robust throughout all specifications of our MRA. While estimated coefficients in case of the former are negative for all five specifications of Table 4, they are only estimated with statistical precision in case of specifications (2) and (3). In case of the dependency ratio, reported coefficients are positive. However, they are only precisely estimated in cases of specifications (1), (3) and (5). Further, this precision is lost if wild-cluster-bootstrapped standard errors are computed.

These results indicate that primary studies accounting for a discretionary impact of ideology are on average associated with higher levels of statistical significance concerning the constraining impact of fiscal rules. In contrast, studies controlling for the dependency ratio yield on average *t*-values in the range of -0.830 to -0.239 implying a loss in statistical significance. More robust evidence is provided by our MRA with respect to the presence of supranational fiscal rules. With the exception of column (4), all coefficients are precisely estimated in our MRA and point to higher levels of statistical significance with respect to

13

the constraining impact of national fiscal rules if primary studies control for the presence of supranational rules.

The results of our MRA also offer insights with respect to the presence of a *publication bias*. Throughout all specifications, obtained coefficients are positive and precisely estimated considering both clustered as well as wild-cluster-bootstrapped standard errors. Relative to the baseline which is defined as publication in a refereed journal, our results indicate that average *t*-values for publications in working papers range between -0.705 and 0.402. These results indicate that on average publications in a refereed journal are more likely to provide evidence for a negative and significant impact of fiscal rules, while results in working papers are more likely to be insignificant.

With respect to properties of the *econometric specification*, our MRA indicates that primary studies running estimations including time fixed effects report stronger negative and significant effects for the impact of fiscal rules. The same holds true in case of bootstrapped errors in comparison to robust standard errors of the baseline effect. However, effects for both variables are not confirmed if wild-cluster-bootstrapped standard errors are considered in our MRA. Furthermore, it has to be emphasised that the obtained result for the baseline effect persists irrespective of whether the observations from primary studies were part of an interaction effect. This observation strengthens our results against the background of the multiple sampling approach employed in our MRA.

The results with respect to *identification strategies* employed in primary studies indicate that accounting for a potential endogeneity matters for the statistical significance with which the efficacy of fiscal rules is measured. In case of column (1), in which variables constituting the identification dummies are employed individually, our results point to lower levels of statistical significance if primary studies control explicitly for preferences or implement quasi-experimental estimation methods. This general observation is confirmed in case of the identification dummies 2, 3, and 4 which comprise of the individual components. Coefficients of all three dummies exhibit positive signs and are, except for identification dummy 4, estimated with statistical precision in our MRA using both clustered as well as wild-cluster-bootstrapped standard errors. Relative to the baseline effect, accounting for a potential endogeneity yields a range of *t*-values between -1.053 and -0.384. Combining these outcomes with the insignificant result for identification dummy 1, the size of the baseline effect in column (2), and the results for preferences as well as quasi-experimental designs in column (1), our MRA provides strong evidence for the presence of potential endogeneity and the relevance of accounting for it when analysing the impact of fiscal rules empirically.

Table 4: Meta-Regression-Analysis: Results using weighted least squares

| | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 | (5) Identification 4 |
|---|---|---|---|---|---|
| *Baseline effect* | −1.874 | −1.254 | −1.762 | −1.758 | −2.057 |
| | [−2.66]** | [−1.39] | [−2.29]** | [−2.24]** | [−2.57]** |
| | {−2.06}** | {−1.01} | {−1.66}* | {−1.62} | {−1.83}* |
| *Classification of dependent variable, Baseline: Primary deficit* | | | | | |
| Secondary deficit | 0.103 | 0.136 | 0.061 | 0.055 | 0.213 |
| | [0.24] | [0.30] | [0.14] | [0.13] | [0.48] |
| | {0.13} | {0.18} | {0.05} | {0.04} | {0.33} |
| Debt | 1.464 | 1.367 | 1.500 | 1.315 | 1.328 |
| | [2.98]*** | [2.73]** | [2.95]*** | [2.61]** | [2.68]** |
| | {2.75}*** | {2.52}** | {2.88}*** | {2.29}** | {2.23}** |
| Primary expenditure | 1.301 | 1.162 | 1.303 | 1.169 | 1.151 |
| | [2.31]** | [2.27]** | [2.23]** | [2.39]** | [2.40]** |
| | {1.41} | {1.52} | {1.37} | {1.59} | {1.61} |
| Secondary expenditure | 1.536 | 1.549 | 1.491 | 1.511 | 1.671 |
| | [3.70]*** | [3.46]*** | [3.50]*** | [3.55]*** | [3.77]*** |
| | {2.88}*** | {2.66}*** | {2.66}*** | {2.88}*** | {3.10}*** |
| Revenue | 2.346 | 2.331 | 2.380 | 2.372 | 2.379 |
| | [3.96]*** | [3.98]*** | [3.82]*** | [3.76]*** | [3.75]*** |
| | {3.10}*** | {3.10}*** | {2.88}*** | {2.75}*** | {3.10}*** |
| *Type of fiscal rule, Baseline: Deficit rule* | | | | | |
| Debt rule | 0.189 | 0.521 | 0.272 | 0.355 | 0.382 |
| | [0.63] | [1.69] | [0.90] | [1.23] | [1.25] |
| | {0.57} | {1.43} | {0.84} | {1.06} | {1.05} |
| Expenditure rule | 0.229 | 0.067 | 0.140 | 0.224 | 0.360 |
| | [0.36] | [0.10] | [0.22] | [0.34] | [0.52] |
| | {0.24} | {0.09} | {0.16} | {0.27} | {0.35} |
| Revenue rule | −0.323 | −0.804 | −0.497 | −0.514 | −0.503 |
| | [−0.45] | [−1.04] | [−0.67] | [−0.70] | [−0.68] |
| | {−0.21} | {−0.63} | {−0.36} | {−0.39} | {−0.35} |
| *Administrative level, Baseline: National* | | | | | |
| State | 0.996 | 1.028 | 1.015 | 0.996 | 0.948 |
| | [1.43] | [1.48] | [1.45] | [1.41] | [1.30] |
| | {1.04} | {1.19} | {1.12} | {1.07} | {0.98} |
| Municipal | −3.029 | −2.053 | −2.626 | −2.662 | −2.969 |
| | [−3.40]*** | [−2.66]** | [−3.46]*** | [−3.45]*** | [−3.36]*** |
| | {−1.93}* | {−1.55} | {−2.08}** | {−1.98}** | {−1.83}* |

Continued on the next page.

## Meta-Regression-Analysis: Results using weighted least squares (continued)

|  | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 | (5) Identification 4 |
|---|---|---|---|---|---|
| *Control variables* | | | | | |
| Ideology | −0.638 | −0.902 | −0.865 | −0.694 | −0.526 |
|  | [−1.41] | [−2.23]** | [−2.19]** | [−1.68] | [−1.14] |
|  | {−1.03} | {−1.79}* | {−1.72}* | {−1.24} | {−0.91} |
| Government stability | −0.011 | 0.503 | 0.235 | −0.030 | −0.024 |
|  | [−0.02] | [1.09] | [0.56] | [−0.06] | [−0.05] |
|  | {−0.08} | {0.87} | {0.39} | {−0.10} | {−0.08} |
| Election years | −0.029 | −0.200 | −0.123 | −0.004 | 0.018 |
|  | [−0.08] | [−0.51] | [−0.34] | [−0.01] | [0.05] |
|  | {−0.07} | {−0.39} | {−0.35} | {−0.01} | {0.05} |
| Debt level | −0.615 | −0.074 | −0.567 | −0.455 | −0.371 |
|  | [−1.13] | [−0.14] | [−1.09] | [−0.89] | [−0.69] |
|  | {−0.67} | {−0.02} | {−0.61} | {−0.49} | {−0.35} |
| Openness | 0.615 | 0.378 | 0.642 | 0.541 | 0.602 |
|  | [1.25] | [0.78] | [1.46] | [1.19] | [1.27] |
|  | {1.04} | {0.65} | {1.28} | {1.01} | {1.12} |
| Dependency ratio | 1.558 | 1.015 | 1.189 | 0.928 | 1.300 |
|  | [2.07]** | [1.52] | [1.78]* | [1.42] | [1.76]* |
|  | {1.58} | {1.39} | {1.58} | {1.31} | {1.43} |
| Supranational rule | −0.960 | −0.834 | −0.873 | −0.556 | −0.728 |
|  | [−2.33]** | [−2.14]** | [−2.45]** | [−1.45] | [−1.95]* |
|  | {−1.98}** | {−1.83}* | {−2.08}** | {−1.30} | {−1.72}* |
| *Publication bias, Baseline: Refereed journal* | | | | | |
| Working paper | 1.474 | 1.656 | 1.527 | 1.357 | 1.352 |
|  | [3.22]*** | [3.53]*** | [3.63]*** | [3.04]*** | [3.10]*** |
|  | {1.96}* | {2.08}** | {2.26}** | {1.85}* | {1.95}* |
| *Econometric specification, Baseline: Robust S.E.* | | | | | |
| Interaction | 0.284 | 0.272 | 0.240 | 0.299 | 0.366 |
|  | [0.57] | [0.54] | [0.47] | [0.59] | [0.74] |
|  | {0.47} | {0.43} | {0.37} | {0.46} | {0.61} |
| Dynamic specification | −0.001 | −0.055 | −0.106 | −0.183 | −0.163 |
|  | [−0.00] | [−0.13] | [−0.27] | [−0.45] | [−0.43] |
|  | {−0.02} | {−0.13} | {−0.19} | {−0.31} | {−0.32} |
| Time fixed effects | −1.565 | −0.965 | −1.490 | −1.344 | −1.303 |
|  | [−2.62]** | [−2.02]* | [−2.63]** | [−2.64]** | [−2.54]** |
|  | {−1.60} | {−1.39} | {−1.65} | {−1.64} | {−1.54} |
| Plain S.E. | 0.084 | −0.352 | −0.043 | −0.205 | −0.032 |
|  | [0.16] | [−0.66] | [−0.08] | [−0.40] | [−0.06] |
|  | {0.11} | {−0.47} | {−0.03} | {−0.28} | {−0.04} |
| Clustered S.E | −0.144 | −0.247 | −0.080 | −0.228 | −0.183 |
|  | [−0.32] | [−0.51] | [−0.18] | [−0.53] | [−0.42] |
|  | {−0.22} | {−0.36} | {−0.11} | {−0.41} | {−0.27} |
| Bootstrapped S.E. | −2.252 | −1.484 | −1.692 | −1.879 | −2.395 |
|  | [−2.17]** | [−1.92]* | [−2.12]** | [−2.29]** | [−2.28]** |
|  | {−1.44} | {−1.61} | {−1.64} | {−1.74}* | {−1.62} |

Meta-Regression-Analysis: Results using weighted least squares (continued)

| | (1)<br>Main specification | (2)<br>Identification 1 | (3)<br>Identification 2 | (4)<br>Identification 3 | (5)<br>Identification 4 |
|---|---|---|---|---|---|
| *Identification* | | | | | |
| Jurisdiction fixed effects | −0.109 | | | | |
| (A) | [−0.24] | | | | |
| | {−0.22} | | | | |
| Preferences and culture | 0.678 | | | | |
| (B) | [1.86]* | | | | |
| | {1.45} | | | | |
| 2SLS *IV* | 0.730 | | | | |
| (C) | [1.59] | | | | |
| | {1.30} | | | | |
| Quasi-experimental | 1.472 | | | | |
| (D) | [1.75]* | | | | |
| | {1.39} | | | | |
| Identification dummy 1 | | −0.802 | | | |
| (A) + (B) + (C) + (D) | | [−1.00] | | | |
| | | {−0.75} | | | |
| Identification dummy 2 | | | 0.709 | | |
| (B) + (C) + (D) | | | [2.11]** | | |
| | | | {1.70}* | | |
| Identification dummy 3 | | | | 1.104 | |
| (C) + (D) | | | | [2.43]** | |
| | | | | {1.91}* | |
| Identification dummy 4 | | | | | 1.673 |
| (D) | | | | | [1.99]* |
| | | | | | {1.52} |
| N | 1033 | 1033 | 1033 | 1033 | 1033 |
| $R^2$ | 0.365 | 0.345 | 0.351 | 0.355 | 0.353 |
| *VIF (maximum)* | 5.453 | 4.887 | 4.506 | 4.508 | 4.701 |

Notes: Clustered *t*-statistics in brackets, wild-cluster-bootstrapped *t*-statistics in braces. All specifications estimated using weighted least squares. Data is trimmed excluding the top and bottom 1% percentiles. Analytical weights equalising the weight for every study are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## Further study dimensions

To gain further insights, we extend our analysis by adding two new variable groups as well as data on journal impact factors. The results for these alterations are summarised in Table 5. To economise on space, we only report the results for the newly added study dimensions are reported. The full set of results is provided by Table A.3 in the appendix.

As adding two new variable groups on *country coverage* and *time coverage* would cause problems with collinearity, it was necessary to drop the information on the *administrative level* in the cases of column (1) and (2). Compared to the estimations in Table 4, the baseline level of statistical significance of the impact of fiscal rules is lower. All four coefficients indicate a statistical significance below the 10% threshold.

Table 5: Meta-Regression-Analysis: Results using further study dimensions

| | (1) Country coverage | (2) Time coverage | (3) Journal | (4) Impact factor |
|---|---|---|---|---|
| *Baseline effect* | −1.271 | 0.494 | −0.400 | −1.336 |
| | [−1.53] | [0.38] | [−0.59] | [−1.55] |
| | {−1.11} | {0.27} | {−0.48} | {−1.12} |
| *Country coverage, Baseline: EU countries* | | | | |
| Switzerland | −1.195 | | | |
| | [−1.72]* | | | |
| | {−1.20} | | | |
| USA | −1.180 | | | |
| | [−0.95] | | | |
| | {−0.59} | | | |
| Euro Members | −0.936 | | | |
| | [−1.15] | | | |
| | {−0.68} | | | |
| OECD Members | 0.277 | | | |
| | [0.55] | | | |
| | {0.46} | | | |
| Other countries | 0.024 | | | |
| | [0.03] | | | |
| | {0.04} | | | |
| *Time coverage* | | | | |
| 1950's | | −3.146 | | |
| | | [−2.02]* | | |
| | | {−1.45} | | |
| 1960's | | 2.424 | | |
| | | [1.97]* | | |
| | | {1.18} | | |
| 1970's | | −2.412 | | |
| | | [−3.12]*** | | |
| | | {−1.55} | | |
| 1980's | | 1.628 | | |
| | | [2.17]** | | |
| | | {1.34} | | |
| 1990's | | −0.761 | | |
| | | [−0.91] | | |
| | | {−0.64} | | |
| 2000's | | −1.506 | | |
| | | [−1.61] | | |
| | | {−0.81} | | |
| 2010's | | −0.846 | | |
| | | [−1.22] | | |
| | | {−0.78} | | |

Continued on the next page.

Meta-Regression-Analysis: Results using further study dimensions (continued)

| | (1)<br>Country coverage | (2)<br>Time coverage | (3)<br>Journal | (4)<br>Impact factor |
|---|---|---|---|---|
| *Publication bias, Baseline in columns: (1-2) Refereed journal; (3-4) Working paper* | | | | |
| Working paper | 0.547<br>[0.78]<br>{0.45} | 1.173<br>[1.92]*<br>{0.82} | | |
| Journal | | | −1.474<br>[−3.22]***<br>{−1.95}* | |
| Impact factor | | | | −0.480<br>[−1.23]<br>{−0.74} |
| *Dropped dimension* | Admin. level | Admin. level | Working paper | Working paper |
| $N$ | 1033 | 1033 | 1033 | 1029 |
| $R^2$ | 0.302 | 0.348 | 0.365 | 0.343 |

Notes: Clustered $t$-statistics in brackets, wild-cluster-bootstrapped $t$-statistics in braces. All specifications estimated using weighted least squares. All estimations exhibit the same set of control variables as the main specification except for the dropped dimension as indicated in the table. The reduced sample size in column (4) is due to lacking information on impact factors. Data is trimmed excluding the top and bottom 1% percentiles. Analytical weights equalising the weight for every study are included.

\* $p < 0.10$, \*\* $p < 0.05$, \*\*\* $p < 0.01$.

Including the *country coverage* as a further dimension asks for the definition of an additional baseline category. In this case, EU countries serve as the baseline. Relative to this subset of countries, only the coverage of Switzerland in primary studies is associated with an additional impact. More precisely, our results indicate that the impact of fiscal rules relative to the baseline is more pronounced in the context of Switzerland with average $t$-values at the 1% level. The picture in case of the time coverage is less clear cut. While primary studies covering data for the 1950's as well as 1970's provide evidence for a stronger statistical impact of fiscal rules well above the 1% threshold, primary studies covering the 1960's as well as 1980's provide evidence for the opposite.

Switching the baseline category from journal publications to working papers in order to test the indicated *publication bias* confirms the results of Table 4. Relative to the baseline, publications in journals are on average associated with negative and statistically significant effects at the 10% level. This is estimated with with statistical precision in our MRA in case of both clustered as well as wild-cluster-bootstrapped standard errors. A differential effect with increasing reputation of the journal by adding information on the impact factor is negated by our MRA. Consequently, a potential publication bias seems to be present for the extensive rather than for the intensive margin. While studies published in refereed journals are generally more likely to report negative and statistical significant effects for the impact of fiscal rules, the level of the statistical significance is not further raised with increasing reputation of the journal.

## Study fixed effects

Table 6 summarises the results for equation (1) with added study fixed effects. In comparison to the main specification and the results of Table 4, alterations in the estimation specification are driven by the lack of within-variation in the respective variables across specifications of studies. As in case of the results for the main specification, the baseline effect indicates a negative and statistically significant impact of fiscal rules. Here, however, the level of statistical significance is clearly above the 1% threshold. Furthermore, the results with respect to the classification of the *dependent variable* are confirmed as well. As before, a weaker impact of fiscal rules is associated with debt, expenditure and revenues.

A different result is obtained in case of the *type of fiscal rules.* In contrast to the main specification, debt rules are now on average associated with a stronger statistical impact compared to deficit rules. However, these effects are only estimated with a statistical precision at the 10% level in our MRA. New results are also obtained with respect to *control variables.* While the results for ideology, the dependency ratio and supranational rules are confirmed, the government stability or openness are associated with a more pronounced impact of fiscal rules. The opposite is true for studies accounting for the impact of elections. Primary studies report on average a reduced constraining impact of fiscal rules when controlling for election years as $t$-values range between -1.048 and -0.409.[15] Results with respect to *identification strategies* are partly confirmed. While the general tendency for a reduction in statistical significance of the effects of fiscal rules are confirmed, these effects are only imprecisely estimated in the context of our MRA.

Table 6: Meta-Regression-Analysis: Results using study fixed effects

| | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 |
|---|---|---|---|---|
| *Baseline effect* | −4.222 | −4.060 | −3.813 | −3.794 |
| | [−5.06]*** | [−5.31]*** | [−5.13]*** | [−5.01]*** |
| *Classification of dependent variable, Baseline: Primary deficit* | | | | |
| Secondary deficit | 0.407 | 0.444 | 0.444 | 0.429 |
| | [0.85] | [0.94] | [0.94] | [0.90] |
| Debt | 1.530 | 1.612 | 1.611 | 1.579 |
| | [5.19]*** | [5.24]*** | [5.21]*** | [5.10]*** |
| Primary expenditure | 0.834 | 0.936 | 0.944 | 0.927 |
| | [1.72]* | [1.94]* | [1.98]* | [1.91]* |
| Secondary expenditure | 1.531 | 1.569 | 1.570 | 1.560 |
| | [2.97]*** | [3.08]*** | [3.09]*** | [3.05]*** |
| Revenue | 2.262 | 2.295 | 2.297 | 2.289 |
| | [4.45]*** | [4.54]*** | [4.54]*** | [4.50]*** |

Continued on the next page.

---

[15]   For evidence for political budget cycles see Rose (2006) and Foremny et al. (2014).

## Meta-Regression-Analysis: Results using study fixed effects (continued)

| | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 |
|---|---|---|---|---|
| *Type of fiscal rule, Baseline: Deficit rule* | | | | |
| Debt rule | −0.481 | −0.519 | −0.522 | −0.524 |
| | [−1.86]* | [−1.97]* | [−1.98]* | [−1.98]* |
| Expenditure rule | −0.273 | −0.271 | −0.268 | −0.261 |
| | [−0.57] | [−0.57] | [−0.57] | [−0.55] |
| Revenue rule | 0.958 | 0.985 | 0.992 | 1.001 |
| | [1.58] | [1.65] | [1.66] | [1.67] |
| *Control variables* | | | | |
| Ideology | −0.124 | −0.128 | −0.131 | −0.088 |
| | [−1.63] | [−1.28] | [−1.11] | [−1.00] |
| Government stability | −2.381 | −2.398 | −2.456 | −2.531 |
| | [−4.16]*** | [−4.33]*** | [−4.55]*** | [−4.45]*** |
| Election years | 3.174 | 3.346 | 3.404 | 3.198 |
| | [3.51]*** | [4.02]*** | [4.07]*** | [3.52]*** |
| Debt level | 0.121 | 0.132 | 0.131 | 0.118 |
| | [0.70] | [0.80] | [0.80] | [0.68] |
| Openness | −0.836 | −0.859 | −0.867 | −0.866 |
| | [−8.61]*** | [−9.37]*** | [−9.43]*** | [−9.43]*** |
| Dependency ratio | 3.888 | 4.035 | 4.043 | 4.007 |
| | [26.88]*** | [40.41]*** | [36.93]*** | [41.58]*** |
| Supranational rule | −0.047 | −0.257 | −0.267 | −0.216 |
| | [−0.23] | [−1.80]* | [−1.71]* | [−1.57] |
| *Econometric specification, Baseline: Robust S.E.* | | | | |
| Interaction | 1.469 | 1.465 | 1.473 | 1.476 |
| | [1.88]* | [1.88]* | [1.89]* | [1.89]* |
| Dynamic specification | −0.898 | −0.982 | −0.978 | −0.899 |
| | [−1.37] | [−1.57] | [−1.58] | [−1.36] |
| Time fixed effects | 0.097 | −0.037 | 0.406 | 0.400 |
| | [0.41] | [−0.08] | [1.69] | [1.81]* |
| Plain S.E. | −0.209 | −0.081 | −0.070 | −0.144 |
| | [−0.70] | [−0.27] | [−0.23] | [−0.48] |
| Bootstrapped S.E. | −0.094 | −0.072 | 0.076 | 0.093 |
| | [−0.25] | [−0.20] | [0.22] | [0.26] |

Continued on the next page.

| | (1)<br>Main specification | (2)<br>Identification 1 | (3)<br>Identification 2 | (4)<br>Identification 3 |
|---|---|---|---|---|
| *Identification* | | | | |
| Jurisdiction fixed effects | 0.827 | | | |
| (A) | [1.47] | | | |
| Preferences and culture | −0.022 | | | |
| (B) | [−0.11] | | | |
| 2SLS *IV* | 0.572 | | | |
| (C) | [1.02] | | | |
| Identification dummy 1 | | 0.584 | | |
| (A) + (B) + (C) + (D) | | [1.47] | | |
| Identification dummy 2 | | | 0.014 | |
| (B) + (C) + (D) | | | [0.05] | |
| Identification dummy 3 | | | | 0.588 |
| (C) + (D) | | | | [1.05] |
| N | 1033 | 1033 | 1033 | 1033 |
| $R^2$ | 0.152 | 0.146 | 0.143 | 0.146 |

Notes: Clustered *t*-statistics in brackets. All specifications estimated using study fixed effects. Alterations in the estimation specification are driven by the lack of within-variation in the respective variables. Data is trimmed excluding the top and bottom 1% percentiles. Analytical weights equalising the weight for every study are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## Ordered probit analysis

Table 7 summarises the results in case of the ordered probit estimation using equation (2) in order to distinguish *t*-values into three categories. Due to the properties of the ordered probit approach, an estimation of a baseline effect is not possible. Furthermore, results for the latent variable are interpreted directly with respect to its sign as well as statistical precision.

As in Table 4, the results of the ordered probit analysis confirm the results for the classification of the *dependent variable* as well as *administrative level* the data of primary studies refer to. While a deviation from the deficit as the dependent variable is associated with a higher likeliness for non-constraining or insignificant results, using municipal data is associated with a higher probability for a statistically significant and constraining effect. Both sets of results are estimated with statistical precision in our MRA. Concerning the set of *control variables*, the main results are again confirmed. Ideology, government stability as well as the presence of supranational rules are associated with a higher likeliness of significant negative results for the impact of fiscal rules. In contrast to the results from the estimation including study fixed effects in Table 6, controlling for openness is found to increase the likeliness for positive or an statistically insignificant impact of rules.

Evidence for a presence of a potential *publication bias* is also provided in the context of the ordered probit estimations. With the exception of column (4) all coefficients are estimated with statistical precision in our MRA and indicate an increased likeliness for

non-existent or positive significant effects if published in a working paper. Also the analysis with respect to *identification strategies* confirms the general tendency of the results of the main specification. While only the coefficient for the identification dummy 3 is estimated with statistical precision, the general tendency of results points to an increased likeliness of insignificant or positive significant results if primary studies account for a potential endogeneity of fiscal rules.

Table 7: Meta-Regression-Analysis: Results using ordered probit

| | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 | (5) Identification 4 |
|---|---|---|---|---|---|
| *Classification of dependent variable, Baseline: Primary deficit* | | | | | |
| Secondary deficit | −0.243 | −0.216 | −0.260 | −0.274 | −0.199 |
| | [−1.08] | [−0.95] | [−1.09] | [−1.13] | [−0.83] |
| Debt | 0.952 | 0.896 | 1.005 | 0.897 | 0.934 |
| | [2.94]*** | [2.89]*** | [3.15]*** | [2.74]*** | [2.81]*** |
| Primary expenditure | 0.806 | 0.770 | 0.829 | 0.770 | 0.766 |
| | [1.96]* | [1.91]* | [1.92]* | [1.97]** | [1.99]** |
| Secondary expenditure | 0.848 | 0.901 | 0.869 | 0.874 | 0.938 |
| | [3.79]*** | [3.97]*** | [4.00]*** | [4.01]*** | [4.14]*** |
| Revenue | 1.513 | 1.521 | 1.547 | 1.541 | 1.540 |
| | [4.85]*** | [5.07]*** | [4.75]*** | [4.70]*** | [4.72]*** |
| *Type of fiscal rule, Baseline: Deficit rule* | | | | | |
| Debt rule | 0.354 | 0.490 | 0.370 | 0.389 | 0.417 |
| | [1.28] | [1.84]* | [1.36] | [1.54] | [1.61] |
| Expenditure rule | 0.355 | 0.264 | 0.325 | 0.368 | 0.412 |
| | [1.34] | [1.01] | [1.22] | [1.39] | [1.44] |
| Revenue rule | −0.485 | −0.692 | −0.550 | −0.525 | −0.550 |
| | [−1.54] | [−2.02]** | [−1.66]* | [−1.68]* | [−1.73]* |
| *Administrative level, Baseline: National* | | | | | |
| State | 0.566 | 0.546 | 0.544 | 0.540 | 0.522 |
| | [1.25] | [1.25] | [1.20] | [1.21] | [1.15] |
| Municipal | −0.984 | −0.659 | −0.966 | −1.023 | −1.092 |
| | [−1.93]* | [−1.65]* | [−2.55]** | [−2.68]*** | [−2.43]** |

Continued on the next page.

23

## Meta-Regression-Analysis: Results using ordered probit (continued)

| | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 | (5) Identification 4 |
|---|---|---|---|---|---|
| *Control variables* | | | | | |
| Ideology | −0.388 | −0.483 | −0.474 | −0.389 | −0.345 |
| | [−1.39] | [−2.10]** | [−1.96]** | [−1.57] | [−1.19] |
| Government stability | −0.647 | −0.326 | −0.459 | −0.623 | −0.549 |
| | [−2.33]** | [−1.43] | [−2.06]** | [−2.39]** | [−1.86]* |
| Election years | 0.219 | 0.108 | 0.143 | 0.218 | 0.202 |
| | [0.93] | [0.47] | [0.61] | [0.94] | [0.88] |
| Debt level | −0.122 | 0.078 | −0.199 | −0.180 | −0.120 |
| | [−0.32] | [0.24] | [−0.61] | [−0.57] | [−0.38] |
| Openness | 0.557 | 0.486 | 0.638 | 0.594 | 0.618 |
| | [1.76]* | [1.57] | [2.14]** | [1.97]** | [1.97]** |
| Dependency ratio | 0.474 | 0.322 | 0.355 | 0.259 | 0.394 |
| | [0.83] | [0.63] | [0.70] | [0.54] | [0.73] |
| Supranational rule | −0.648 | −0.603 | −0.559 | −0.419 | −0.501 |
| | [−1.84]* | [−2.21]** | [−2.18]** | [−1.75]* | [−2.12]** |
| *Publication bias, Baseline: Refereed journal* | | | | | |
| Working paper | 0.423 | 0.520 | 0.433 | 0.339 | 0.362 |
| | [1.79]* | [2.14]** | [1.94]* | [1.56] | [1.70]* |
| *Econometric specification, Baseline: Robust S.E.* | | | | | |
| Interaction | 0.139 | 0.123 | 0.127 | 0.149 | 0.180 |
| | [0.50] | [0.44] | [0.45] | [0.53] | [0.67] |
| Dynamic specification | −0.142 | −0.168 | −0.240 | −0.263 | −0.266 |
| | [−0.48] | [−0.56] | [−0.85] | [−0.92] | [−0.99] |
| Time fixed effects | −0.436 | −0.235 | −0.525 | −0.486 | −0.447 |
| | [−1.13] | [−0.78] | [−1.63] | [−1.61] | [−1.50] |
| Plain S.E. | −0.086 | −0.232 | −0.075 | −0.138 | −0.072 |
| | [−0.27] | [−0.75] | [−0.24] | [−0.43] | [−0.22] |
| Clustered S.E. | −0.632 | −0.640 | −0.543 | −0.621 | −0.586 |
| | [−2.14]** | [−2.06]** | [−1.87]* | [−2.19]** | [−2.04]** |
| Bootstrapped S.E. | −1.019 | −0.804 | −0.877 | −0.993 | −1.137 |
| | [−1.55] | [−1.55] | [−1.57] | [−1.78]* | [−1.64] |

Continued on the next page.

| | (1) Main specification | (2) Identification 1 | (3) Identification 2 | (4) Identification 3 | (5) Identification 4 |
|---|---|---|---|---|---|
| *Identification* | | | | | |
| Jurisdiction fixed effects | −0.286 | | | | |
| (A) | [−0.73] | | | | |
| Preferences and culture | 0.118 | | | | |
| (B) | [0.64] | | | | |
| 2SLS *IV* | 0.542 | | | | |
| (C) | [1.50] | | | | |
| Quasi-experimental | 0.495 | | | | |
| (D) | [0.91] | | | | |
| Identification dummy 1 | | −0.567 | | | |
| (A) + (B) + (C) + (D) | | [−1.13] | | | |
| Identification dummy 2 | | | 0.298 | | |
| (B) + (C) + (D) | | | [1.51] | | |
| Identification dummy 3 | | | | 0.599 | |
| (C) + (D) | | | | [2.01]** | |
| Identification dummy 4 | | | | | 0.670 |
| (D) | | | | | [1.33] |
| *Cut-offs* | | | | | |
| Cut 1 | −0.359 | −0.632 | −0.282 | −0.289 | −0.165 |
| *Insignificant* | [−0.87] | [−1.26] | [−0.65] | [−0.66] | [−0.38] |
| Cut 2 | 2.150 | 1.858 | 2.203 | 2.203 | 2.318 |
| *Positive significant* | [4.94]*** | [3.65]*** | [4.94]*** | [4.95]*** | [5.07]*** |
| N | 1033 | 1033 | 1033 | 1033 | 1033 |

Notes: Clustered *t*-statistics in brackets. All specifications estimated using ordered probit. Outcomes coded as -1 for negative significant, 0 for insignificant and +1 for positive significant. Data is trimmed excluding the top and bottom 1% percentiles. Analytical weights equalising the weight for every study are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

## Effect-size analysis

Table 8 summarises the results with respect to the effect size for the largest homogeneous sub-group of observations. This sub-group consists of studies testing for the impact of fiscal rules on both the primary and secondary deficit expressed relative to the GDP and measured in terms of a dummy variable. As a consequence, sample size is reduced significantly including only nine primary studies and results should only be interpreted with caution and seen as suggestive. The alterations in the estimation specification are due to the reduced sample size and the resulting lack in coverage of particular dimensions as well as collinearity.

Against this background, the baseline effect is estimated to range between -1.5 and 1.3 and can be interpreted as a reduction in the deficit in percent of the GDP if a deficit rule is present. With the exception of columns (3) and (4), these effects are estimated with statistical precision. The effect size is roughly reduced by 0.5 percentage points if the impact on the secondary rather than the primary deficit is considered. For the sub-group

of these nine studies the efficacy of fiscal rules at the state level is strongly reduced and even found to be deficit increasing compared to the national level.

While not very robust, results for *control variables* such as ideology are confirmed. However, in case of variables such as supranational rules, signs of obtained coefficients switch. The same holds true for effects with respect to *country coverage* and *time coverage*. For instance, in Euro member states, fiscal rules are found to be associated with increased deficits. Results with respect to the *time coverage* are again not clear cut. While studies using data for the 1960s and 2010s are associated with stronger constraining effects, this is not the case for the 1970s. Results with respect to a potential *publication bias* are also confirmed by the analysis for the effect size. Observations obtained from journal publications are associated with a larger constraining effect from fiscal rules compared to observation from working papers. The difference amounts up to 2.5 percentage points of GDP. This effect is increasing with the reputation of the journal measured by its impact factor.

Table 8: Meta-Regression-Analysis: Results for the effect size

| | (1) Main specification | (2) Country coverage | (3) Time coverage | (4) Journal | (5) Impact factor |
|---|---|---|---|---|---|
| *Baseline effect* | −1.231 | −1.306 | −1.357 | 1.280 | −1.521 |
| | [−2.53]** | [−3.23]** | [−1.34] | [0.73] | [−3.06]** |
| | {−1.32} | {−1.48} | {−0.72} | {0.62} | {−1.50} |
| *Classification of dependent variable, Baseline: Primary deficit* | | | | | |
| Secondary deficit | 0.546 | 0.546 | 0.546 | 0.546 | 0.471 |
| | [3.13]** | [3.11]** | [3.11]** | [3.13]** | [2.06]* |
| | {7.1e+18}*** | {7.1e+18}*** | {7.1e+18}*** | {7.1e+18}*** | {0.71} |
| *Type of fiscal rule, Baseline: Deficit rule* | | | | | |
| Debt rule | 0.332 | 0.332 | 0.325 | 0.332 | 0.280 |
| | [1.72] | [1.70] | [1.76] | [1.72] | [1.61] |
| | {0.91} | {0.91} | {0.95} | {0.91} | {1.16} |
| Expenditure rule | 2.822 | 2.822 | 2.821 | 2.822 | 2.802 |
| | [1.87]* | [1.86] | [1.86] | [1.87]* | [1.86] |
| | {0.76} | {0.76} | {0.76} | {0.76} | {0.76} |
| Revenue rule | −3.148 | −3.147 | −3.139 | −3.148 | −2.839 |
| | [−2.06]* | [−2.04]* | [−2.03]* | [−2.06]* | [−1.97]* |
| | {−0.75} | {−0.75} | {−0.75} | {−0.75} | {−0.75} |
| *Administrative level, Baseline: National* | | | | | |
| State | 4.122 | | | 4.122 | 3.484 |
| | [3.17]** | | | [3.17]** | [2.89]** |
| | {1.55} | | | {1.55} | {1.40} |

Continued on the next page.

| | (1) Main specification | (2) Country coverage | (3) Time coverage | (4) Journal | (5) Impact factor |
|---|---|---|---|---|---|
| *Control variables* | | | | | |
| Ideology | −2.544 | −0.448 | 1.047 | −2.544 | −3.075 |
| | [−4.23]*** | [−0.66] | [0.97] | [−4.23]*** | [−3.15]** |
| | {−2.09}** | {−0.72} | {0.66} | {−2.09}** | {−1.56} |
| Government stability | −0.295 | 1.743 | 3.492 | −0.295 | −0.064 |
| | [−0.48] | [3.68]*** | [5.01]*** | [−0.48] | [−0.10] |
| | {−0.45} | {0.87} | {1.55} | {−0.45} | {−0.14} |
| Election years | −0.112 | −0.125 | −0.112 | −0.112 | −0.016 |
| | [−0.64] | [−0.71] | [−0.64] | [−0.64] | [−0.07] |
| | {−0.62} | {−0.62} | {−0.62} | {−0.62} | {−0.04} |
| Debt level | −0.061 | −0.061 | −0.060 | −0.061 | 0.293 |
| | [−1.71] | [−1.70] | [−1.72] | [−1.71] | [0.61] |
| | {−0.95} | {−0.95} | {−0.95} | {−0.95} | {0.71} |
| Supranational rule | 1.791 | −0.245 | −1.951 | 1.791 | 1.700 |
| | [2.50]** | [−0.31] | [−1.78] | [2.50]** | [1.45] |
| | {1.74}* | {−0.68} | {−1.06} | {1.74}* | {1.04} |
| *Econometric specification, Baseline: Robust S.E.* | | | | | |
| Interaction | 0.887 | 0.886 | 0.859 | 0.887 | 0.064 |
| | [1.03] | [1.01] | [0.82] | [1.03] | [0.11] |
| | {0.68} | {0.68} | {0.68} | {0.68} | {0.11} |
| Dynamic specification | −0.596 | −0.596 | −0.596 | −0.596 | −0.569 |
| | [−4.11]*** | [−4.08]*** | [−4.08]*** | [−4.11]*** | [−2.74]** |
| | {−3.14}*** | {−3.14}*** | {−3.14}*** | {−3.14}*** | {−0.83} |
| Time fixed effects | 0.859 | 0.856 | 0.291 | 0.859 | −0.540 |
| | [0.67] | [0.65] | [0.21] | [0.67] | [−1.01] |
| | {0.65} | {0.59} | {0.33} | {0.65} | {−0.84} |
| Plain S.E. | −1.374 | −1.333 | −4.603 | −1.374 | −0.478 |
| | [−1.28] | [−1.17] | [−2.40]** | [−1.28] | [−0.93] |
| | {−0.72} | {−0.68} | {−1.55} | {−0.72} | {−0.57} |
| Clustered S.E. | −0.726 | −0.686 | 0.538 | −0.726 | 3.621 |
| | [−0.83] | [−0.77] | [0.27] | [−0.83] | [1.76] |
| | {−0.88} | {−0.90} | {0.31} | {−0.88} | {1.76}* |
| Bootstrapped S.E. | −0.616 | −0.613 | −0.540 | −0.616 | 0.231 |
| | [−2.60]** | [−2.55]** | [−1.19] | [−2.60]** | [0.53] |
| | {−1.12} | {−1.12} | {−0.77} | {−1.12} | {0.42} |
| *Identification* | | | | | |
| Jurisdiction fixed effects | −0.664 | −0.657 | −0.002 | −0.664 | 1.498 |
| | [−0.43] | [−0.41] | [−0.00] | [−0.43] | [1.92]* |
| | {−0.33} | {−0.33} | {−0.01} | {−0.33} | {1.65} |

Continued on the next page.

Meta-Regression-Analysis: Results for the effect size (continued)

| | (1)<br>Main specification | (2)<br>Country coverage | (3)<br>Time coverage | (4)<br>Journal | (5)<br>Impact factor |
|---|---|---|---|---|---|
| *Country coverage, Baseline: EU countries* | | | | | |
| Euro Members | | 2.064<br>[3.30]**<br>{1.42} | | | |
| OECD Members | | 0.030<br>[0.16]<br>{0.27} | | | |
| Other countries | | 0.068<br>[0.41]<br>{0.63} | | | |
| *Time coverage* | | | | | |
| 1960's | | | −4.019<br>[−1.94]*<br>{−1.41} | | |
| 1970's | | | 3.288<br>[2.77]**<br>{1.04} | | |
| 2010's | | | −0.505<br>[−2.24]*<br>{−0.68} | | |
| *Publication bias, Baseline in columns: (1-3) Refereed journal; (4-5) Working paper* | | | | | |
| Working paper | 2.512<br>[1.98]*<br>{0.77} | 2.523<br>[1.96]*<br>{0.77} | 2.478<br>[1.70]<br>{0.72} | | |
| Journal | | | | −2.512<br>[−1.98]*<br>{−0.76} | |
| Impact factor | | | | | −2.112<br>[−1.96]*<br>{−1.67}* |
| $N$ | 176 | 176 | 176 | 176 | 176 |
| $R^2$ | 0.376 | 0.376 | 0.377 | 0.376 | 0.372 |

Notes: Clustered $t$-statistics in brackets, wild-cluster-bootstrapped $t$-statistics in braces. All specifications estimated using weighted least squares. The changes in the estimation specification are due to the reduced sample size and the resulting lack in coverage of particular dimensions as well as collinearity. All remaining variables are backed by more than one study. Data is trimmed excluding the top and bottom 1% percentiles. Analytical weights equalising the weight for every study are included.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

# 6 Conclusions

The empirical literature on the budgetary implications of numerical fiscal rules has been growing quickly over the recent years and has reached a critical mass rendering MRA feasible and appropriate to summarise the overall evidence. Our paper is the first MRA in this particular field of literature which explores and summarises existing empirical evidence from 30 studies from the last decade in a systematic way. In this regard, we provide a consensus estimate with respect to the sign as well as statistical significance with which fiscal rules influence fiscal aggregates. Furthermore, for a small homogeneous subgroup of studies, we also provide a consensus estimate with respect to the size of effects. However, our MRA contributes to this field of literature beyond a simple stocktaking. First, it provides guidance to the crucial discussion with respect to the identification of causal effects within this field of literature. It puts special emphasis on the impact resulting from the use of different types of identification strategies. Second, our analysis provides insights with respect to the role of included control variables and guides the determination of an appropriate specification for future analyses of the budgetary impact of fiscal rules.

Our MRA points to a statistically significant and constraining impact of fiscal rules on fiscal aggregates at the national level. Obtained results indicate a level of significance around the 10% threshold. These results hold particularly strong for deficits but less so for debt, expenditures or revenues. Relative to the baseline of deficit rules, alternative types of fiscal rules do not exert a differential impact. Concerning the effect size, albeit only measurable for a small subset of studies, the results show that deficit rules on average reduce the primary deficit between 1.2 and 1.5 percent of GDP.

With respect to the identification of causal effects, our analysis provides supportive evidence that fiscal rules must not be treated as exogenous. Our results show that coefficients tend to lose their statistical significance with increasing refinement of applied identification strategies. While the inclusion of jurisdiction fixed effects as a rough measure for fiscal preferences does not make a systematic difference, the employment of instrumental variables or quasi-experimental designs lead to markedly lower levels of significance and a less constraining impact. In light of these insights, our results emphasise the necessity for using appropriate identification strategies in future research in order to advance our knowledge on the causal impact of numerical fiscal rules.

A further insight emerges from our meta-regressions-analysis with respect to minimum standards for future studies and necessary controls. Our results indicate that controlling for government stability, political budget cycles, demographic characteristics as well as the presence of a supranational rule is relevant for the measured impact of fiscal rules. Consequently, the set of control variables should involve proxies for these dimensions whenever possible. With respect to differences across federal levels, our results indicate that constraining effects of fiscal rules are the strongest for studies using municipal data.

Furthermore, our MRA provides evidence for a publication bias. Results obtained from working papers are on average associated with lower levels of statistical significance compared to those from journal articles.

Summarising, the existing empirical evidence points to constraining effects of rules on fiscal aggregates. However, results have to be interpreted carefully as on the one hand our analysis points to a bias if the potential endogeneity is not taken explicitly into account and on the other hand provides evidence for the presence of a publication bias. Against this background, it will be crucial to develop and implement new convincing identification strategies in order to provide robust and causally interpretable empirical evidence in this field of literature.

# References

Abrams, B. and W. Dougan (1986). The effects of constitutional restraints on governmental spending. *Public Choice 49*(2), 101–116.

Afonso, A. and S. Hauptmeier (2009). Fiscal behaviour in the European Union: Rules, fiscal decentralization and government indebtedness. *ECB Working Paper 1054 (May)*.

Angrist, J. D. and J. Pischke (2009). *Mostly harmless econometrics*. Princeton and Oxford: Princeton University Press.

Argimón, I. and P. Hernández de Cos (2012). Fiscal rules and federalism as determinants of budget performance: An empirical investigation for the Spanish case. *Public Finance Review 40(1)*, 30–65.

Bails, D. (1982). A critique on the effectiveness of tax-expenditure limitations. *Public Choice 38*(2), 129–138.

Bartolini, D. and R. Santolini (2009). Fiscal rules and the opportunistic behaviour of the incumbent politician: Evidence from Italian municipalities. *CESifo Working Paper 2605 (April)*.

Baskaran, T., L. P. Feld, and J. Schnellenbach (2014). Fiscal federalism, decentralization and economic growth: Survey and meta-analysis. *CESifo Working Paper No. 4985 (September)*.

Benito, B., F. Bastida, and C. Vicente (2013). Creating room for manoeuvre: A strategy to generate political budget cycles under fiscal rules. *Kyklos 66*(4), 467–496.

Brooks, L., Y. Halberstam, and J. Phillips (2012). Spending within limits: Evidence from municipal fiscal restraints. *Finance and Economics Discussion Series, Federal Reserve Board 2012-52 (April)*.

Burret, H. T. and L. P. Feld (2014). Political institutions and fiscal policies: A note on budget rules and fiscal federalism. *CESifo DICE Report 1/2014 (March)*.

Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics 90*(3), 414–427.

Cameron, A. C. and D. L. Miller (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources 50*(2), 1–53.

Cameron, A. C. and P. K. Trivedi (2010). *Microeconometrics using Stata*. College Station, Texas: Stata Press.

Card, D., J. Kluve, and A. Weber (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal 120*(548), F452–F477.

Cevik, S. and K. Teksoz (2014). Deep roots of fiscal behaviour. *IMF Working Paper 14-45 (March)*.

Christofzik, D. I. and S. G. Kessing (2014). Does fiscal oversight matter? *CESifo Working Paper No. 5023 (October)*.

Coutinho, L., D. Georgiou, M. Heracleous, A. Michaelides, and S. Tsani (2013). Limiting fiscal procyclicality: Evidence from resource-rich countries. *CEPR Discussion Paper No. 9672 (October)*.

Dahan, M. and M. Strawczynski (2013). Fiscal rules and the composition of government expenditures in OECD countries. *Journal of Policy Analysis and Management 32(3)*, 484–504.

Debrun, X. and M. S. Kumar (2007). The discipline-enhancing role of fiscal institutions: Theory and empirical evidence. *IMF Working Paper 07-171 (July)*.

Debrun, X., L. Moulin, A. Turrini, J. Ayuso-i Casals, and M. S. Kumar (2008). Tied to the mast? National fiscal rules in the European Union. *Economic Policy 23*(4), 297–362.

Ebeke, C. and D. Ölçer (2013). Fiscal policy over the election cycle in low-income countries. *IMF Working Paper 13-153 (June)*.

Escolano, J., L. Eyraud, M. M. Badia, J. Sarnes, and A. Tuladhar (2012). Fiscal performance, institutional design and decentralization in European Union countries. *IMF Working Paper 12-45 (February)*.

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

Foremny, D., R. Freier, M.-D. Moessinger, and M. Yeter (2014). Overlapping political budget cycles. *ZEW Discussion Paper No. 14-099 (December)*.

Gollwitzer, S. (2010). Budget institutions and fiscal performance in Africa. *Journal of African Economies 20*(1), 111–152.

Greene, W. H. (2003). *Econometric analysis.* Upper Saddle River: Prentice Hall.

Grembi, V., T. Nannicini, and U. Troiano (2012). Policy responses to fiscal restraints: A difference-in-discontinuities design. *CESifo Working Paper No. 3999 (November)*.

Hallerberg, M., R. Strauch, and J. von Hagen (2007). The design of fiscal rules and forms

of governance in European Union countries. *European Journal of Political Economy 23*, 338–359.

Harbord, R. M. and J. P. T. Higgins (2008). Meta-regression in Stata. *The Stata Journal 8*(4), 493–519.

Hedbávný, P., O. Schneider, and J. Zápal (2007). A fiscal rule that has teeth: A suggestion for a 'fiscal sustainability council' underpinned by the financial markets. *Czech Economic Review 1*(1), 32–53.

Heinemann, F., S. Osterloh, and A. Kalb (2014). Sovereign risk premia: The link between fiscal rules and stability culture. *Journal of International Money and Finance 41*, 110–127.

Imbeau, L. M. (2004). Public deficits and surpluses in federated states: A review of the public choice empirical literature. *Journal of Public Finance and Public Choice 22*(3), 123–145.

Kirchgässner, G. (2002). The effects of fiscal institutions on public finance: A survey of the empirical evidence. In S. L. Winer and H. Shibata (Eds.), *Political Economy and Public Finance: The Role of Political Economy in Theory and Practice of Public Economics*, pp. 145–177.

Klomp, J. and J. de Haan (2010). Inflation and central bank independence: A meta-regression analysis. *Journal of Economic Surveys 24*(4), 593–621.

Kopits, G. and S. Symansky (1998). Fiscal policy rules. *IMF Occasional Paper 162*.

Krogstrup, S. and S. Wälti (2008). Do fiscal rules cause budgetary outcomes? *Public Choice 136*(1-2), 123–138.

Luechinger, S. and C. A. Schaltegger (2013). Fiscal rules, budget deficits and budget projections. *International Tax and Public Finance 20*(5), 785–807.

Marneffe, W., B. van Aarle, W. van der Wielen, and L. Vereeck (2010). The impact of fiscal rules on public finances: Theory and empirical evidence for the euro area. *CESifo Working Paper No. 3303 (December)*.

Miyazaki, T. (2014). Fiscal reform and fiscal sustainability: Evidence from Australia and Sweden. *International Review of Economics and Finance 33*, 141–151.

Nelson, J. P. and P. E. Kenedy (2009). The use (and abuse) of meta-analysis in environmental and natural resource economics: An assessment. *Environmental and Resource Economics 42*(3), 345–377.

Nerlich, C. and W. H. Reuter (2013). The design of national fiscal frameworks and their budgetary impact. *ECB Working Paper No. 1588 (September)*.

Neyapti, B. (2013). Fiscal decentralization, fiscal rules and fiscal discipline. *Economics Letters 121*, 528–532.

Poterba, J. M. (1996). Budget institutions and fiscal policy in the U.S. states. *American Economic Review 86*(2), 395–400.

Primo, D. M. (2006). Stop us before we spend again: Institutional constraints on government spending. *Economics & Politics 18(3)*, 269–312.

Rose, S. (2006). Do fiscal rules dampen the political business cycle? *Public Choice 128*(3-4), 407–431.

Schaltegger, C. A. and L. P. Feld (2009). Do large cabinets favor large governments? Evidence on the fiscal commons problem for Swiss cantons. *Journal of Public Economics 93*, 35–47.

Shadbegian, R. J. (1996). Do tax and expenditure limitations affect the size and growth of state government? *Contemporary Economic Policy 14(1)*, 22–35.

Stanley, T., H. Doucouliagos, M. Giles, J. H. Heckemeyer, R. J. Johnston, P. Laroche, J. P. Nelson, M. Paldam, J. Poot, and G. Pugh (2013). Meta-analysis of economics research reporting guidelines. *Journal of Economic Surveys 27(2)*, 390–394.

Stanley, T. D. and H. Doucouliagos (2012). *Meta-regression analysis in economics and business*. Routledge Advances in Research Methods. London: Routledge.

Tapp, S. (2013). The use and effectivenes of fiscal rules in Canadian provinces. *CanadianPublic Policy 39(1)*, 45–70.

Tapsoba, R. (2012). Do national numerical fiscal rules really shape fiscal behaviours in developing countries? A treatment effect evaluation. *Economic Modelling 29*, 1356–1369.

Thornton, J. (2009). On the relation between central and sub-national government fiscal balances. *Applied Economics Letters 16*(6), 603–608.

von Hagen, J. (2002). Fiscal rules, fiscal institutions, and fiscal performance. *The Economic and Social Review 33(3)*, 263–284.

# A   Appendix

Table A.1: Overview of tables considered in the meta-regression-analysis

| # | Authors and Year | Keyword | Published | Journal/Series |
|---|---|---|---|---|
| 1 | Afonso and Hauptmeier (2009) | Fiscal rule | No | ECB Working Paper Series |
| 2 | Argimón and Hernández de Cos (2012) | Fiscal rule | Yes | Public Finance Review |
| 3 | Bartolini and Santolini (2009) | Fiscal rule | No | CESifo Working Paper Series |
| 4 | Benito et al. (2013) | Fiscal rule | Yes | Kyklos |
| 5 | Brooks et al. (2012) | Fiscal restraint | No | Finance and Economics Discussion Series |
| 6 | Cevik and Teksoz (2014) | Fiscal rule | No | IMF Working Paper |
| 7 | Christofzik and Kessing (2014) | Fiscal rule | No | CESifo Working Paper Series |
| 8 | Coutinho et al. (2013) | Fiscal rule | No | CEPR Discussion Paper Series |
| 9 | Dahan and Strawczynski (2013) | Fiscal rule | Yes | Journal of Policy Analysis and Management |
| 10 | Debrun and Kumar (2007) | Fiscal rule | No | IMF Working Paper Series |
| 11 | Debrun et al. (2008) | Fiscal rule | Yes | Economic Policy |
| 12 | Ebeke and Ölçer (2013) | Fiscal rule | No | IMF Working Paper Series |
| 13 | Escolano et al. (2012) | Fiscal rule | No | IMF Working Paper Series |
| 14 | Foremny (2014) | Fiscal rule | Yes | European Journal of Political Economy |
| 15 | Gollwitzer (2010) | Budget rule | Yes | Journal of African Economies |
| 16 | Grembi et al. (2012) | Fiscal rule | No | CESifo Working Paper Series |
| 17 | Hallerberg et al. (2007) | Fiscal rule | Yes | European Journal of Political Economy |
| 18 | Hedbávný et al. (2007) | Fiscal rule | Yes | Czech Economic Review |
| 19 | Krogstrup and Wälti (2008) | Fiscal rule | Yes | Public Choice |
| 20 | Luechinger and Schaltegger (2013) | Fiscal rule | Yes | International Tax and Public Finance |
| 21 | Marneffe et al. (2010) | Fiscal rule | No | CESifo Working Paper Series |
| 22 | Miyazaki (2014) | Fiscal rule | Yes | International Review of Economics and Finance |
| 23 | Nerlich and Reuter (2013) | Fiscal rule | No | ECB Working Paper Series |
| 24 | Neyapti (2013) | Fiscal rule | Yes | Economics Letters |
| 25 | Primo (2006) | Budget rule | Yes | Economics & Politics |
| 26 | Rose (2006) | Fiscal rule | Yes | Public Choice |
| 27 | Schaltegger and Feld (2009) | Fiscal rule | Yes | Journal of Public Economics |
| 28 | Tapp (2013) | Fiscal rule | Yes | Canadian Public Policy |
| 29 | Tapsoba (2012) | Fiscal rule | Yes | Economic Modelling |
| 30 | Thornton (2009) | Fiscal rule | Yes | Applied Economics Letters |

*Notes*: Result of the search on *EconLit* based on the six keywords: *fiscal rule*, *fiscal restraint*, *debt brake*, *debt rule*, *budget rule*, and *deficit rule*. If an individual paper has been found using several keyword, we mention the first with a positive search result.

Table A.2: Variable Description

| Variable | Description |
|----------|-------------|
| | *Dummy = 1 if the primary study's* |

*Classification of dependent variable*

| | |
|---|---|
| Debt | Dependent variable is debt. |
| Secondary deficit | Dependent variable is secondary deficit. |
| Primary deficit | Dependent variable is primary deficit. |
| Secondary expenditure | Dependent variable is secondary expenditure, including subcategories of spending, e.g., expenditure for interest payments. |
| Primary expenditure | Dependent variable is primary expenditure. |
| Revenue | Dependent variable is revenue, inlcuding subcategories, e.g., tax revenue or social security contributions. |

*Type of fiscal rule*

| | |
|---|---|
| Deficit rule | Fiscal rule indicator comprises a deficit rule. |
| Debt rule | Fiscal rule indicator comprises a debt rule. |
| Expenditure rule | Fiscal rule indicator comprises an expenditure rule. |
| Revenue rule | Fiscal rule indicator comprises a revenue rule. |

*Country coverage*

| | |
|---|---|
| Switzerland | Country coverage comprises Switzerland. |
| USA | Country coverage comprises the USA. |
| European Union | Country coverage comprises states that belong to the European Union. The dummy is only equal to one if the full set of countries belongs to the European Union. |
| Euro area | Country coverage comprises states that belong to the Euro area. The dummy is only equal to one if the full set of countries belongs to the Euro area. |
| OECD | Country coverage comprises states that belong to the OECD. The dummy is only equal to one if the full set of countries belongs to the OECD. |
| Other country | Country coverage does not include Switzerland, the USA, European Union contries, Euro area countries, or OECD countries. |

*Administrative level*

| | |
|---|---|
| National | Administrative level is the national level. |
| State | Administrative level is the state level. |
| Municipality | Administrative level is the municipal level. |

*Control variables*

| | |
|---|---|
| Ideology | Control variables comprise government ideology indicators. |
| Government stability | Control variables comprise government stability indicators, e.g., tenure of the government or specific index values. |
| Election year | Control variables comprise variables indicating election years. |
| Debt | Control variables comprise debt. |
| Preferences and culture | Control variables comprise preferences and cultural proxies, e.g., language dummies, regional dummies, or fiscal preferences dummies. |
| Dependency ratio | Control variables comprise dependency ratio indicators, e.g., share of working population in total population. |
| Output gap | Control variables comprise the output gap. |
| Growth | Control variables comprise GDP growth indicators. |
| Openness | Control variables comprise openness indicators, e.g., sum of exports and imports to GDP. |
| Budgetary rules | Control variables comprise indicators for budgetary rules, e.g., delegation vs. contract approach. |
| Supranational rule | Control variables comprise an indicator for a supranational fiscal rule, e.g., the Stability and Growth Pact. |

Continued on next page.

Table A.2: Variable Description (continued)

| Variable | Description<br>*Dummy = 1 if the primary study's* |
|---|---|
| *Publication bias* | |
| Refereed journal | Observations are published in a refereed journal. |
| Working paper | Observations are published in a working paper. |
| *Econometric specification* | |
| Dynamic specification | Econometric specification includes the lagged dependent variable. |
| Time fixed effects | Econometric specification includes time fixed effects. |
| Plain S.E. | Econometric specification includes plain standard errors. |
| Robust S.E. | Econometric specification includes robust standard errors. |
| Clustered S.E. | Econometric specification includes clustered standard errors. |
| Bootstrapped S.E. | Econometric specification includes bootstrapped standard errors. |
| *Identification* | |
| Jurisd. fixed effects | Econometric specification includes jurisdiction fixed effects, e.g., country fixed effects. |
| Preferences and culture | Control variables comprise preferences and cultural proxies, e.g., language dummies, regional dummies, or fiscal preferences dummies. |
| 2SLS (IV) | Estimation method is an instrumental variables approach (2SLS) with the fiscal rule indicator as the instrumented variable. |
| Quasi-experimental | Estimation method relies on quasi-experimental methods, e.g., matching. |

Figure A.1: Boxplot for the classification of the dependent variable based on the full sample without windsorising
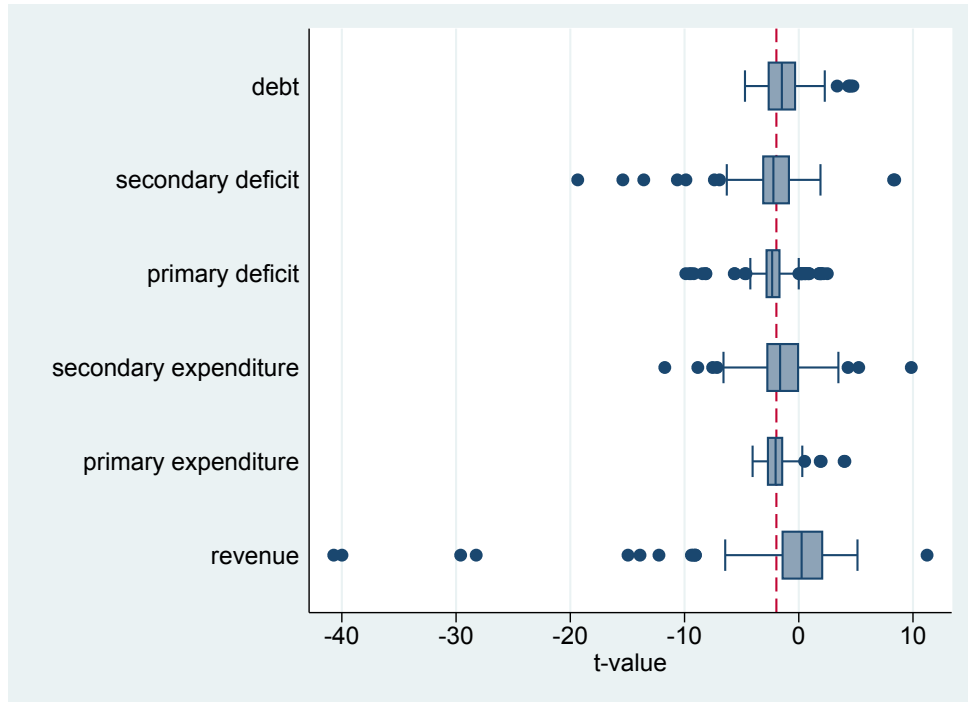


Notes: Data of the full sample without trimming of the top and bottom 1% percentiles and without analytical weights.

## Table A.3: Meta-Regression-Analysis: Results using further study dimensions

| | (1)<br>Country coverage | (2)<br>Time coverage | (3)<br>Journal | (4)<br>Impact factor |
|---|---|---|---|---|
| *Baseline effect* | −1.271 | 0.494 | −0.400 | −1.336 |
| | [−1.53] | [0.38] | [−0.59] | [−1.55] |
| | {−1.11} | {0.27} | {−0.48} | {−1.12} |
| *Classification of dependent variable, Baseline: Primary deficit* | | | | |
| Secondary deficit | 0.454 | 0.142 | 0.103 | −0.182 |
| | [0.92] | [0.32] | [0.24] | [−0.40] |
| | {0.67} | {0.27} | {0.13} | {−0.37} |
| Debt | 1.646 | 0.819 | 1.464 | 1.126 |
| | [2.96]*** | [1.90]* | [2.98]*** | [2.19]** |
| | {2.52}** | {1.55} | {2.75}*** | {1.76}* |
| Primary expenditure | 1.310 | 0.938 | 1.301 | 1.248 |
| | [2.46]** | [1.60] | [2.31]** | [2.17]** |
| | {1.80}* | {1.07} | {1.41} | {1.41} |
| Secondary expenditure | 1.989 | 1.417 | 1.536 | 1.687 |
| | [2.87]*** | [3.35]*** | [3.70]*** | [3.73]*** |
| | {1.88}* | {2.41}** | {2.88}*** | {3.10}*** |
| Revenue | 2.705 | 2.231 | 2.346 | 2.392 |
| | [3.88]*** | [3.81]*** | [3.96]*** | [3.89]*** |
| | {2.75}*** | {2.52}** | {3.10}*** | {3.1e+19}*** |
| *Type of fiscal rule, Baseline: Deficit rule* | | | | |
| Debt rule | 0.766 | 0.312 | 0.189 | 0.462 |
| | [1.55] | [0.78] | [0.63] | [1.27] |
| | {1.09} | {0.66} | {0.57} | {1.02} |
| Expenditure rule | 0.247 | −0.027 | 0.229 | 0.386 |
| | [0.36] | [−0.04] | [0.36] | [0.54] |
| | {0.26} | {−0.00} | {0.24} | {0.37} |
| Revenue rule | −0.167 | −0.791 | −0.323 | −0.428 |
| | [−0.20] | [−0.94] | [−0.45] | [−0.50] |
| | {−0.15} | {−0.54} | {−0.21} | {−0.22} |
| *Administrative level, Baseline: National* | | | | |
| State | | | 0.996 | 0.265 |
| | | | [1.43] | [0.42] |
| | | | {1.04} | {0.35} |
| Municipal | | | −3.029 | −2.495 |
| | | | [−3.40]*** | [−2.37]** |
| | | | {−1.93}* | {−1.27} |

Continued on the next page.

## Meta-Regression-Analysis: Results using further study dimensions (continued)

| | (1)<br>Country coverage | (2)<br>Time coverage | (3)<br>Journal | (4)<br>Impact factor |
|---|---|---|---|---|
| *Control variables* | | | | |
| Ideology | −0.388 | −1.299 | −0.638 | −0.350 |
| | [−0.73] | [−2.67]** | [−1.41] | [−0.75] |
| | {−0.44} | {−1.52} | {−1.03} | {−0.51} |
| Government stability | −0.827 | 0.155 | −0.011 | −0.489 |
| | [−1.01] | [0.27] | [−0.02] | [−1.05] |
| | {−0.78} | {0.18} | {−0.08} | {−0.85} |
| Election years | −0.225 | −0.140 | −0.029 | 0.058 |
| | [−0.39] | [−0.33] | [−0.08] | [0.15] |
| | {−0.21} | {−0.23} | {−0.07} | {0.15} |
| Debt level | 0.626 | 1.274 | −0.615 | −0.079 |
| | [1.07] | [2.03]* | [−1.13] | [−0.16] |
| | {0.88} | {1.30} | {−0.67} | {−0.02} |
| Openness | −0.433 | 0.845 | 0.615 | 0.516 |
| | [−0.60] | [1.13] | [1.25] | [1.01] |
| | {−0.50} | {0.83} | {1.04} | {0.81} |
| Dependency ratio | 0.508 | 2.102 | 1.558 | 1.525 |
| | [0.48] | [2.78]*** | [2.07]** | [1.86]* |
| | {0.39} | {2.12}** | {1.58} | {1.36} |
| Supranational rule | −1.350 | −1.365 | −0.960 | −1.229 |
| | [−2.32]** | [−2.62]** | [−2.33]** | [−2.84]*** |
| | {−2.00}** | {−1.73}* | {−1.98}** | {−2.37}** |
| *Econometric specification, Baseline: Robust S.E.* | | | | |
| Interaction | −0.060 | −0.095 | 0.284 | 0.215 |
| | [−0.11] | [−0.18] | [0.57] | [0.43] |
| | {−0.10} | {−0.12} | {0.47} | {0.36} |
| Dynamic specification | −0.167 | −0.238 | −0.001 | 0.180 |
| | [−0.36] | [−0.52] | [−0.00] | [0.38] |
| | {−0.25} | {−0.36} | {−0.02} | {0.22} |
| Time fixed effects | −0.340 | −0.304 | −1.565 | −0.969 |
| | [−0.64] | [−0.53] | [−2.62]** | [−1.90]* |
| | {−0.46} | {−0.33} | {−1.60} | {−1.30} |
| Plain S.E. | −1.026 | −0.941 | 0.084 | −0.144 |
| | [−1.49] | [−1.43] | [0.16] | [−0.24] |
| | {−1.01} | {−0.97} | {0.11} | {−0.15} |
| Clustered S.E. | −1.161 | −1.887 | −0.144 | −0.192 |
| | [−1.82]* | [−3.33]*** | [−0.32] | [−0.49] |
| | {−1.37} | {−2.66}*** | {−0.22} | {−0.38} |
| Bootstrapped S.E. | −0.662 | −2.253 | −2.252 | −2.419 |
| | [−0.49] | [−2.28]** | [−2.17]** | [−2.31]** |
| | {−0.34} | {−1.62} | {−1.44} | {−1.58} |

Continued on the next page.

Meta-Regression-Analysis: Results using further study dimensions (continued)

| | (1)<br>Country coverage | (2)<br>Time coverage | (3)<br>Journal | (4)<br>Impact factor |
|---|---|---|---|---|
| *Identification* | | | | |
| Jurisdiction fixed effects | −0.674 | −0.156 | −0.109 | −0.282 |
| | [−1.36] | [−0.22] | [−0.24] | [−0.54] |
| | {−0.99} | {−0.23} | {−0.22} | {−0.46} |
| Preferences and culture | 0.234 | 0.113 | 0.678 | 0.518 |
| | [1.02] | [0.43] | [1.86]* | [1.41] |
| | {0.93} | {0.34} | {1.45} | {1.16} |
| 2SLS *IV* | 0.826 | 0.581 | 0.730 | 0.771 |
| | [1.87]* | [1.39] | [1.59] | [1.64] |
| | {1.56} | {1.16} | {1.30} | {1.29} |
| Quasi-experimental | −0.027 | −0.204 | 1.472 | 1.652 |
| | [−0.03] | [−0.26] | [1.75]* | [1.69] |
| | {−0.05} | {−0.21} | {1.39} | {1.25} |
| *Country coverage, Baseline: EU countries* | | | | |
| Switzerland | −1.195 | | | |
| | [−1.72]* | | | |
| | {−1.20} | | | |
| USA | −1.180 | | | |
| | [−0.95] | | | |
| | {−0.59} | | | |
| Euro Members | −0.936 | | | |
| | [−1.15] | | | |
| | {−0.68} | | | |
| OECD Members | 0.277 | | | |
| | [0.55] | | | |
| | {0.46} | | | |
| Other countries | 0.024 | | | |
| | [0.03] | | | |
| | {0.04} | | | |

Continued on the next page.

## Meta-Regression-Analysis: Results using further study dimensions (continued)

| | (1)<br>Country coverage | (2)<br>Time coverage | (3)<br>Journal | (4)<br>Impact factor |
|---|---|---|---|---|
| *Time coverage* | | | | |
| 1950's | | −3.146 | | |
| | | [−2.02]* | | |
| | | {−1.45} | | |
| 1960's | | 2.424 | | |
| | | [1.97]* | | |
| | | {1.18} | | |
| 1970's | | −2.412 | | |
| | | [−3.12]*** | | |
| | | {−1.55} | | |
| 1980's | | 1.628 | | |
| | | [2.17]** | | |
| | | {1.34} | | |
| 1990's | | −0.761 | | |
| | | [−0.91] | | |
| | | {−0.64} | | |
| 2000's | | −1.506 | | |
| | | [−1.61] | | |
| | | {−0.81} | | |
| 2010's | | −0.846 | | |
| | | [−1.22] | | |
| | | {−0.78} | | |
| *Publication bias, Baseline in columns: (1-2) Refereed journal; (3-4) Working paper* | | | | |
| Working paper | 0.547 | 1.173 | | |
| | [0.78] | [1.92]* | | |
| | {0.45} | {0.82} | | |
| Journal | | | −1.474 | |
| | | | [−3.22]*** | |
| | | | {−1.95}* | |
| Impact factor | | | | −0.480 |
| | | | | [−1.23] |
| | | | | {−0.74} |
| $N$ | 1033 | 1033 | 1033 | 1029 |
| $R^2$ | 0.302 | 0.348 | 0.365 | 0.343 |

Notes: Clustered $t$-statistics in brackets, wild-cluster-bootstrapped $t$-statistics in braces. Reduced sample size in column (4) is due to lacking information on impact factors. All specifications estimated using weighted least squares. Data is trimmed excluding the top and bottom 1% percentiles. Analytical weights equalising the weight for every study are included.

\* $p < 0.10$, \*\* $p < 0.05$, \*\*\* $p < 0.01$.