

Conselhos fiscais independentes ajudam a melhorar o desempenho fiscal?*

FERNANDO COVELLI BENELLI[†]

FABIANA ROCHA[‡]

VITOR KAYO DE OLIVEIRA[§]

Sumário

1. Introdução	410
2. Razões para usar o método de controle sintético	413
3. Amostra	415
4. Resultados	418
5. Testes adicionais	425
6. Conclusões	427
Apêndice A.	431

Palavras-chave

conselho fiscal, controle sintético, reformas institucionais

JEL Codes

H62, C33, D02



Resumo • Abstract

O objetivo deste artigo é avaliar o impacto sobre o resultado fiscal dos conselhos fiscais independentes. Sua principal contribuição consiste na aplicação da metodologia de controle sintético, mais rigorosa quanto ao controle de possíveis vieses nas estimações. Os resultados encontrados apontam para a ausência de qualquer influência dos conselhos fiscais sobre o resultado fiscal, tanto ao nível individual dos países como nas avaliações de grupo, e independente da qualidade institucional do órgão. Diversos testes de robustez confirmam os achados do estudo.

1. Introdução

A estrutura institucional que governa a execução da política fiscal desempenha um papel importante para a qualidade das finanças públicas. Desta forma, regras e conselhos fiscais bem estabelecidos e dotados de credibilidade são fundamentais. Estes podem evitar que o processo político seja marcado por vieses de gastos e déficits que acabem resultando em um setor público grande e ineficiente, podem garantir a

*Fabiana Rocha agradece o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq. Fernando Benelli agradece pela bolsa de auxílio à pesquisa do Programa de Pós-Graduação (PPG), patrocinado pelo Banco Central do Brasil. As opiniões expressas no artigo são de responsabilidade dos autores e não refletem necessariamente os do Banco Central do Brasil. Quaisquer erros são de responsabilidade exclusiva dos autores.

[†]Banco Central do Brasil. Avenida Paulista 1804, Bela Vista, São Paulo, SP, CEP 01310-922, Brasil. fernando.benelli@bcb.gov.br [0000-0003-4686-5897](https://orcid.org/0000-0003-4686-5897)

[‡]Universidade de São Paulo, Faculdade de Economia, Administração e Contabilidade (FEA/USP). Avenida Professor Luciano Gualberto 908, Butantã, São Paulo, SP, CEP 05508-010, Brasil. frocha@usp.br [0000-0001-7192-3353](https://orcid.org/0000-0001-7192-3353)

[§]Mestre em Economia, IPE-USP. vitor.kayo.de.oliveira@gmail.com

fernando.benelli@bcb.gov.br frocha@usp.br vitor.kayo.de.oliveira@gmail.com

estabilidade da política fiscal e podem ainda reduzir o escopo para atividades de *rent seeking*.

Regras fiscais podem ser definidas como um *target* (outargets) para a política fiscal, assim como um conjunto de condições que devem ser seguidas caso esse *target* (ou *targets*) não seja(m) alcançado(s). São inúmeros os estudos que buscam avaliar o impacto das regras sobre o comportamento fiscal do governo, tendo a maioria encontrado evidência de que estas resultam em menores gastos e maiores saldos fiscais.¹

Conselhos fiscais (CFs), por sua vez, são organismos públicos independentes (não partidários) e técnicos que têm como objetivo preparar previsões macroeconômicas para o orçamento, monitorar o desempenho fiscal e/ou aconselhar o governo em assuntos fiscais.²

Os conselhos fiscais, ao contrário das regras fiscais, foram em sua maioria estabelecidos recentemente e, conseqüentemente, foram menos avaliados.^{3,4}

Debrun e Kumar (2007) encontram evidência de que instituições fiscais (regras e agências não partidárias) afetam positivamente a disciplina fiscal atuando, assim, como estímulos para o comprometimento dos *policymakers*. Entretanto, eles verificam também a presença de causalidade reversa, o que implica que são os governos fiscalmente responsáveis que adotam regras e conselhos fiscais independentes a fim de revelar a natureza de suas preferências (sinalização).⁵ Debrun, Gérard, e Harris (2011) estimam um efeito fraco dos conselhos fiscais sobre a mídia e sobre as mudanças fiscais. Não encontram, contudo, um efeito direto dos conselhos sobre o resultado fiscal. Nerlich e Reuter (2013) apontam que regras fiscais ajudam a melhorar o saldo fiscal e o impacto é mais forte se as regras fiscais são combinadas com um conselho fiscal independente e uma estrutura orçamentária efetiva de médio prazo.⁶ Coletta e Infantino (2015) concluem que os conselhos fiscais têm um impacto positivo sobre o desempenho fiscal e que o status legal do conselho fiscal (se é formalmente atrelado ao Parlamento ou ao governo, sua forma de financiamento e o acesso à

¹Por exemplo, Von Hagen (1992); Von Hagen e Harden (1995); Poterba (1996); Eichengreen, Hausmann, e Von Hagen (1999); Hallerberg, Strauch, e Von Hagen (2009); Tapsoba (2012); Heinemann, Osterloh, e Kalb (2014); Reuter (2015); Badinger e Reuter (2017).

²Pelo seu papel em estimular a escolha de políticas fiscalmente responsáveis através do aumento do custo de reputação e político de escolhas inadequadas, os conselhos fiscais são chamados de *fiscal watchdogs*.

³Por exemplo, Debrun, Hauner, e Kumar (2009) e Calmfors e Wren-Lewis (2011).

⁴No final de 2016 foi criado no Brasil um conselho desta natureza, denominado Instituição Fiscal Independente (IFI).

⁵A tentativa de uma abordagem econométrica mais rigorosa fica comprometida pelo fato de os resultados serem sensíveis à escolha das variáveis instrumentais.

⁶É interessante que o impacto das regras fiscais depende de sua especificação. Regras de orçamento equilibrado afetam quase qualquer categoria de gasto, enquanto regras de dívida têm efeitos importantes somente sobre categorias específicas.

informação interna) é sua característica mais importante. De acordo com [Debrun e Kinda \(2017\)](#), países com regras fiscais melhor desenhadas apresentam melhor desempenho fiscal, mas a mera existência de conselhos fiscais não conduz a saldos fiscais mais elevados. Resultados positivos só aparecem quando são consideradas certas características dos conselhos fiscais. Conselhos mais independentes, que monitoram regras fiscais numéricas, que contribuem tecnicamente (por exemplo, avaliando as previsões oficiais) e que têm grande impacto na mídia tendem a exibir melhor desempenho fiscal. Conselhos fiscais e suas características também estão associados a menores erros de previsão do saldo primário e do crescimento real. Contudo, como observam [Debrun e Kinda \(2017\)](#), esses resultados devem ser vistos como correlações condicionais robustas devido às dificuldades envolvidas na estimação. [Beetsma et al. \(2018\)](#) fornecem evidências de que conselhos fiscais parecem eliminar os vieses otimistas que caracterizam as previsões orçamentárias, além de reduzirem os erros de previsão. Além disso, estimulam a adesão a regras de orçamento equilibrado e de gastos.

O objetivo deste artigo é analisar o efeito potencial dos conselhos fiscais sobre o resultado fiscal, medido pelo saldo estrutural como porcentagem do produto, usando o método de controle sintético ([Abadie, Diamond, & Hainmueller, 2010](#); [Abadie & Gardeazabal, 2003](#)). Procura, assim, estimar uma causalidade que é particularmente difícil de estabelecer, como mostra a literatura resumida anteriormente. O artigo contribui para a relativamente escassa literatura que usa controle sintético para avaliar o impacto de *policy* ([Billmeier & Nannicini, 2013](#); [Abadie, Diamond, & Hainmueller, 2015](#)) e parece ser a primeira contribuição à avaliação de impacto dos conselhos fiscais usando esta metodologia. A amostra é composta por 20 países tratados (que adotaram conselhos fiscais) e 36 países não tratados durante o período 2000 a 2017.

O método de controle sintético procura aproximar, o melhor possível, a trajetória de uma variável de resultado para uma única unidade de interesse (unidade tratada) no período pré-intervenção com a trajetória de um grupo de controle. O grupo de controle consiste numa média ponderada de um número positivo de unidades, o que explica ser denominado de “sintético”. A trajetória da variável de resultado no período pós-intervenção corresponderia então à evolução da variável de resultado da unidade tratada caso a intervenção não tivesse ocorrido, sendo o estimador de controle sintético dado pela diferença entre as duas trajetórias no período depois da intervenção.

O artigo está organizado em 5 seções além desta introdução. A [seção 2](#) justifica a escolha do método de controle sintético a partir de sua comparação com o método de diferença-em-diferenças. A [seção 3](#) discute a amostra, em particular as diferentes características dos conselhos fiscais que formam o grupo de tratamento. A [seção 4](#) apresenta inicialmente os resultados da estimação de diferença-em-diferenças e em seguida os resultados da estimação de controle sintético, tanto para cada país

individualmente quanto para o conjunto dos países. A seção 5 mostra os resultados de alguns testes de robustez. A seção 6 resume as principais conclusões e traz sugestões para pesquisa futura.

2. Razões para usar o método de controle sintético⁷

A estimação de diferença-em-diferenças (DID) aparece como primeira alternativa para isolar o impacto da adoção de conselhos fiscais de outras influências que afetam o saldo estrutural. Ele envolve a identificação de um tratamento específico (no caso, a adoção de CFs) e compara a diferença nos resultados fiscais antes e depois do tratamento para os países tratados com a diferença nos resultados antes e depois da adoção do tratamento para os países não tratados.

O problema é que a escolha do grupo de comparação que mimetiza o contrafactual é feita pelo próprio pesquisador. Com isso, a escolha de países para compor o grupo de controle seria feita de forma *ad hoc*, o que levaria a dúvidas de se eles constituem um contrafactual adequado.

O método de controle sintético reduz a ambiguidade na escolha do grupo de controle, pois usa um processo baseado nos dados para criar um controle artificial que é dado pela média ponderada de várias unidades de um conjunto de unidades potenciais de comparação (*donor pool*). O método torna explícita a contribuição de cada unidade de comparação para o contrafactual, sendo que algumas unidades potenciais não comporão o controle sintético, uma vez que receberão peso zero.^{8,9} No

⁷Para mais detalhes da comparação entre o método de controle sintético e os métodos de regressão, ver Abadie et al. (2015).

⁸O procedimento envolve os seguintes passos. Primeiro, achar os pesos iniciais (*V-weights*) usando regressão. Achar os pesos ótimos (*W-weights*). O vetor W é escolhido de tal forma a aproximar o melhor possível o vetor X_i , um vetor ($K \times 1$) das variáveis preditoras para a unidade de tratamento no período pré-intervenção e a matriz X_J , uma matriz ($K \times J$) contendo as mesmas K variáveis preditoras para o mesmo período mas para todas as possíveis unidades de controles no *donor pool*. O vetor W é, assim, escolhido para minimizar a diferença

$$\|X_i - X_J W\|_V = \sqrt{(X_i - X_J W)' V (X_i - X_J W)}.$$

O objetivo final é minimizar o erro quadrático médio de previsão (*root mean square prediction error* – RMSPE) da variável de resultado no período pré-intervenção. Então, são encontrados *V-weights* alternativos via método de iteração e calculados novos *W-weights* até que o RMSPE seja minimizado. O RMSPE é dado por

$$RMSPE = \sqrt{\frac{1}{T_0} \sum_1^{T_0} \left(Y_{it} - \sum_{j=2}^{J+1} w_j^* Y_j^t \right)^2}.$$

⁹Ao tornar explícita a contribuição de cada unidade de comparação para o contrafactual, o método de controle sintético permite ainda combinar o uso de análise quantitativa e qualitativa. Os pesos, contudo, raramente são reportados.

DID tradicional, por sua vez, o contrafactual é construído usando a média simples de todas as unidades de controle (condicional nas covariadas).

Por construção, o caso sintético se parece com o caso tratado antes do tratamento. Como o desempenho da unidade sintética antes do evento é similar ao da unidade sendo avaliada, a estimativa de impacto é dada simplesmente pela diferença nos valores da variável de resultado para a unidade tratada e a sintética depois do evento.

Enquanto a estimação DID, na maioria das vezes, fornece somente estimativas do efeito médio, o método de controle sintético permite que sejam obtidas estimativas individuais para cada caso. A obtenção de estimativas para cada unidade tratada (país) de uma vez faz com que seja possível explorar a heterogeneidade dos efeitos da adoção dos CFs de forma mais flexível. Uma vez que o desenho de cada instituição fiscal independente, assim como o ambiente em que ela atua, afeta o seu impacto, assumir heterogeneidade é fundamental. Além disso, a média ponderada de países similares fornece um contrafactual mais adequado do que a média simples de todos os países de controle.

O método de controle sintético pode reduzir o problema causado pela omissão de variáveis. Desde que o pareamento pré-intervenção seja adequado (a unidade sintética é tal que ela lembra muito a trajetória pré-tratamento da variável de resultado da unidade tratada), [Abadie et al. \(2010\)](#) provam que o viés causado por *confounders* não observados variantes no tempo tende a zero à medida que o período pré-intervenção cresce. Intuitivamente a explicação é que trajetórias semelhantes da variável de resultado para períodos longos somente poderiam ser produzidas se as unidades (países) tivessem preditores, tanto observados quanto não observados, da variável de resultado muito similares e se o efeito destes preditores sobre a variável de resultado também fosse muito semelhante.

Finalmente, o método de controle sintético permite o uso de experimentos de placebo para garantir inferências válidas na presença de erros correlacionados, o que não ocorre com a estimação DID. Como mostram [Bertrand, Duflo, e Mullainathan \(2004\)](#), os desvios-padrão dos estimadores em artigos que usam a estimação DID com dados de muitos anos estão subestimados devido à presença de correlação nos erros.

Obviamente, o método de controle sintético também tem suas limitações. Ele não leva em conta outras fontes de endogeneidade como a causalidade reversa. Esta é uma limitação importante, uma vez que a adoção de conselhos fiscais independentes é pelo menos em parte resultante de perspectivas negativas de desempenho fiscal.

O grupo de controle ideal é composto de unidades que não sofreram nenhum choque idiossincrático importante capaz de afetar a variável de resultado de interesse no período amostral analisado, em especial no período pós-tratamento. O problema é que não é possível excluir do *donor pool* países que tomaram outras medidas, que não a adoção de conselhos fiscais, para melhorar o seu resultado fiscal. Ainda que estas medidas sejam menores, o resultado acumulado pode ser significativo.

Contudo, neste caso, caso um efeito seja encontrado, as estimativas provavelmente estão subestimando o verdadeiro impacto dos conselhos fiscais independentes.

3. Amostra

A variável de interesse é o resultado estrutural (SB), como proporção do PIB potencial, em termos anuais, elaborado pela Divisão de Assuntos Fiscais do FMI e publicado no relatório World Economic Outlook (WEO) de outubro de 2018.

De acordo com o Fiscal Council Dataset, elaborado pelo Departamento de Assuntos Fiscais do FMI, 39 países haviam adotado CFs até 2016. Os conselhos fiscais da Croácia, Eslovênia e Japão não foram considerados em razão de controvérsias quanto à sua atuação como instituições independentes. Quatro conselhos foram instaurados antes da data de início do período em análise e assim também foram desconsiderados: Bélgica (1989 e 1994), Dinamarca (1962), Estados Unidos (1974) e México (2000). Três conselhos adotados entre 2000 e 2009, Coreia do Sul (2003), Suécia (2007) e Canadá (2008), não foram considerados por apresentarem um período pré-tratamento muito curto.¹⁰ Como o Fiscal Council Dataset apontou que o conselho da Hungria havia sido significativamente enfraquecido, a Hungria foi incluída no grupo de controle, ou seja, como um país que não possui conselho fiscal. Além disso, devido à forte crise fiscal sofrida pela Grécia, esta foi excluída da amostra. A movimentação anormal de recursos do governo com credores externos prejudicaria uma análise da efetiva contribuição de seu conselho instituído em 2010.

Desta forma, o nosso grupo de tratamento é composto por 20 países e nosso grupo de controle por 36 países. O período analisado é de 2000 até 2017. A [Tabela 1](#) apresenta os países que compõem o grupo de tratamento e o ano em que foi implementado o conselho fiscal.¹¹

A [Figura 1](#) mostra que os países tratados exibiam, na média, saldos estruturais mais baixos que os países de controle antes e durante a crise fiscal de 2008–2009. Essa situação se reverte após 2012, período em que a maioria dos CFs foi criada.

Tanto modelos teóricos como avaliações empíricas¹² indicam que o desenho ideal de um CF, no que diz respeito à sua capacidade de influenciar a trajetória da

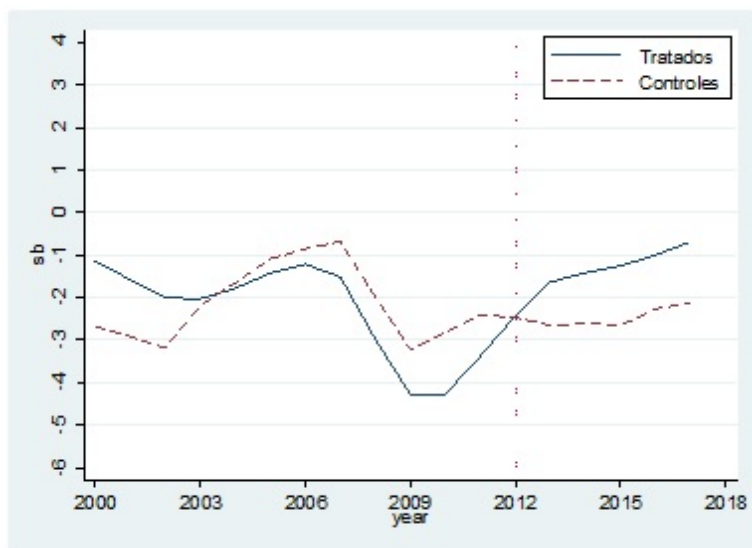
¹⁰Por decisão *ad hoc*, a extensão mínima de ajuste pré-tratamento foi estipulada em 10 anos. Testes realizados com valores inferiores a esse limite (disponíveis diretamente com os autores), aumentam demasiadamente a RMSPE, levando a uma sobrestimação dos efeitos ajustados e a uma diminuição do poder do teste. Extensões superiores a 10 anos restringem o número de tratados, deixando de fora países importantes como Alemanha e Reino Unido, entre outros.

¹¹Os países que formam o grupo de controle são: Argentina, Bósnia e Herzegovina, Brasil, Bulgária, Cazaquistão, China, Croácia, República Checa, República Dominicana, Equador, Granada, Guiana, Hong Kong, Hungria, Islândia, Índia, Indonésia, Israel, Japão, Malásia, Marrocos, Nova Zelândia, Noruega, Paraguai, Filipinas, Polônia, Rússia, Cingapura, Eslovênia, Suriname, Suíça, Tailândia, Tunísia, Turquia, Ucrânia e Uruguai.

¹²Por exemplo, [Debrun e Kumar \(2008\)](#); [Debrun e Kinda \(2017\)](#); [Beetsma et al. \(2018\)](#).

Tabela 1. Países tratados e anos de implementação dos conselhos fiscais

Alemanha	2010	África do Sul	2014
Reino Unido	2010	Chile	2014
Eslováquia	2011	Chipre	2014
Colômbia	2012	Espanha	2014
Irlanda	2012	Estônia	2014
Portugal	2012	Holanda	2014
Austrália	2013	Itália	2014
Áustria	2013	Luxemburgo	2014
Finlândia	2013	Malta	2015
França	2013	Peru	2015

**Figura 1.** Trajetórias das médias aritméticas simples dos resultados primários estruturais/PIB para os países tratados e não tratados

dívida pública, deve atender aos cinco atributos essenciais listados abaixo, de acordo com o Fiscal Council Dataset. Assim, foram classificados como CFs fortes apenas os países cujos conselhos atendiam simultaneamente a todos esses atributos e como CFs fracos aqueles para os quais um ou mais deles estavam ausentes.

1. *Elaborar ou avaliar previsões macroeconômicas*: Previsões oficiais da arrecadação fiscal são comumente utilizadas na elaboração do orçamento. Com base nelas é calculado o volume total de despesas para os exercícios seguintes. Sendo assim, previsões demasiadamente otimistas podem ser produzidas no intuito de escapar das restrições impostas pelas regras fiscais e ampliar o espaço para mais gasto público.¹³
2. *Monitorar Regras Fiscais*: Na grande maioria dos casos, a implementação do CF ocorreu em países que já possuíam regras fiscais pré-estabelecidas. A existência de metas numéricas para o resultado fiscal constitui uma clara referência para os trabalhos do conselho, sugerindo a complementaridade entre os dois mecanismos de controle fiscal.¹⁴
3. *Impacto na mídia*: Diferentemente dos bancos centrais, os CFs não têm poderes para influenciar diretamente a política fiscal. Sua influência provém principalmente da participação no debate público e dos impactos reputacionais e eleitorais de suas avaliações.¹⁵
4. *Participar no processo orçamentário*: Um conselho fiscal é considerado participativo quando suas previsões são empregadas no orçamento; quando estas exercem poderes restritivos, ou seja, quando resultam em explicações quanto ao desvio dos valores preditos; quando há obrigação formal do governo e/ou do congresso em consultar o conselho e quando o conselho tem algum poder de interromper o processo orçamentário. Participar da elaboração da lei orçamentária amplia os impactos reputacionais das análises do conselho, reduzindo o viés deficitário e a prociclicidade que muitas vezes impactam a política fiscal discricionária.

¹³Frankel e Schreger (2012) encontram evidência de que a participação de instituições fiscais na avaliação e/ou elaboração de previsões reduz o viés otimista. Jonung e Larch (2006) acham que previsões elaboradas por uma instituição independente são menos sujeitas a influência política do que as produzidas pelo Ministério da Fazenda

¹⁴Nerlich e Reuter (2013) encontram que os impactos positivos das regras sobre o resultado primário são ampliados quando monitoradas por instituições independentes. Coletta e Infantino (2015) restringem a influência benéfica do monitoramento apenas às regras com forte base legal.

¹⁵Debrun et al. (2011) mostram que a presença dos CFs na mídia costuma ser mais intensa em tempos de instabilidade econômica/fiscal, justamente quando sua mensagem é mais necessária. Além disso, Debrun e Kinda (2017) argumentam que esta presença é capaz de influir tanto na trajetória do resultado primário como na redução dos erros de previsão macroeconômicas.

5. Independência: a independência é considerada uma condição necessária ao bom funcionamento do CF, qualquer que seja seu desenho e objetivos, pois sem ela o conselho perde sua credibilidade, invalidando seu papel de prestador de recomendações qualificadas e isentas à sociedade.¹⁶

A [Tabela 2](#) mostra a classificação dos conselhos fiscais dos 20 países que compõem o grupo de tratamento entre fortes e fracos e destaca, para os países com CFs fracos, quais características não são observadas.

Cabem alguns comentários com relação aos conselhos avaliados. A Alemanha possui dois conselhos fiscais, o German Council of Economic Experts e o Independent Advisory Board to the German Stability Council, criados, respectivamente em 1963 e 2010. Suas características foram agrupadas como se fosse um único conselho, o mesmo ocorrendo com o conselho da Holanda, o Raad von State, de 2014, que foi precedido pelo Netherlands Bureau for Economic Policy Analysis, de 1945. O documento de criação do conselho fiscal da Austrália (Parliamentary Budget Office), estabelecido em 2012, foi precedido pelo Charter of Budget Honesty Act, de 1998, um documento legal que impôs à elaboração do orçamento procedimentos semelhantes ao de uma instituição independente. Também agrupamos as disposições contidas em ambos os documentos.

4. Resultados

4.1 Resultados: diferença-em-diferenças

Inicialmente são apresentados os resultados da estimação do modelo de diferenças-em-diferenças (DID) tradicional:

¹⁶A independência envolve vários aspectos. i) Não partidarismo, sustentado pela existência de garantias legais e mecanismos que apoiem a independência operacional da instituição. Neste último, avalia-se o isolamento das opiniões do conselho de qualquer interferência partidária. Um conselho legalmente constituído como uma entidade separada do Poder a que está vinculado e com mandatos bem definidos atende automaticamente aos dois critérios. Quando o conselho é parte integral do congresso ou de algum ministério, mesmo havendo independência legal, a operacional dependerá da reputação de não-partidarismo no processo orçamentário e no debate público. ii) Salvaguardas no orçamento que ocorrem quando os recursos orçamentários destinados ao conselho são protegidos de possíveis atos políticos retaliatórios. As salvaguardas são consideradas existentes se o orçamento for definido pelo banco central, se parte do orçamento geral vier do legislativo, ou seja, for protegido das decisões do executivo, se for garantido por dotações orçamentárias em itens destacados no orçamento, ou estiver sujeito a qualquer outra garantia geralmente concedida a instituições independentes, como as agências reguladoras. iii) Presença de pessoal qualificado, ou seja, incluir entre seus membros acadêmicos e/ou especialistas em política fiscal. O conselho não deve incluir de forma nenhuma apenas políticos ou servidores públicos. iv) Indicação idônea dos membros. Embora o mais comum é que seja realizado pelo Executivo, um crescente número de países tem optado em eleger os membros do conselho pelo Legislativo, o que pode ser visto como uma prevenção contra interferências do governo no mandato dos membros eleitos. Em alguns casos raros, as nomeações são feitas pelo chefe de Estado, ou por instituições independente, como o banco central ou algum escritório de auditoria.

Tabela 2. Classificação dos CFs em Fortes e Fracos, segundo as características presentes

	Previsões	Monitoramento de Regras Fiscais	Impacto na Mídia	Participação no Processo Orçamentário	Independência
Fortes					
Alemanha	✓	✓	✓	✓	✓
Austrália	✓	✓	✓	✓	✓
Áustria	✓	✓	✓	✓	✓
Chipre	✓	✓	✓	✓	✓
Estônia	✓	✓	✓	✓	✓
Finlândia	✓	✓	✓	✓	✓
França	✓	✓	✓	✓	✓
Holanda	✓	✓	✓	✓	✓
Itália	✓	✓	✓	✓	✓
Portugal	✓	✓	✓	✓	✓
Reino Unido	✓	✓	✓	✓	✓
Fracos					
África do Sul					
Chile	✓			✓	
Colômbia					
Eslováquia		✓	✓		✓
Espanha	✓	✓	✓	✓	
Irlanda	✓	✓	✓		✓
Luxemburgo	✓		✓	✓	
Malta	✓	✓	✓		
Peru	✓	✓	✓		

$$Y_{it} = \alpha + X'_{it}\pi + \rho CF_{it} + \gamma_i + \lambda_t + \theta_i t + \varepsilon_{it}, \tag{1}$$

onde Y_{it} é o resultado primário estrutural do país i no ano t ; CF_{it} é uma variável *dummy* que assume valor igual a 1 nos anos de atuação dos conselhos fiscais; X_{it} é uma matriz contendo as características observáveis dos países (ver seção A.1); γ_i e λ_t são os efeitos fixos de país e tempo (ano), respectivamente; θ_i é a tendência linear específica ao país; e α é o intercepto geral da regressão. Supondo $E[\varepsilon_{it} | X_{it}, CF_{it}, \gamma_i, \lambda_t, \theta_i] = 0$, coeficiente ρ é o parâmetro de interesse, pois captura o efeito do tratamento sobre os tratados. Os resultados são apresentados na [Tabela 3](#).

Tabela 3. Efeitos dos conselhos fiscais sobre o resultado estrutural (DID tradicional)

Variável Dependente	Saldo Estrutural (%PIB)		
	Todos os Conselhos	Conselhos Fortes	Conselhos Fracos
CF	1,581*** (0,526)	2,240*** (0,648)	1,210* (0,696)
Observações	886	724	689
Número de Tratados	20	11	9
Número de Controles	36	36	36

Notas: Erros-padrão em parênteses. Termos constantes, covariadas, tendências lineares específicas aos países, efeitos fixos de ano e país incluídos, mas não reportados. Erros-padrão clusterizados por país. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

A adoção de conselhos fiscais independentes gera uma elevação média de 1,58pp no superávit primário estrutural. Os conselhos fiscais fortes geram um efeito ainda maior. Países que adotam conselhos fiscais fortes apresentam superávits 2,24pp maiores do que aqueles que não adotam conselhos fiscais. No caso dos conselhos fracos, o efeito permanece positivo, mas é bem menor em magnitude (1,2pp).

4.2 Resultados: controle sintético

Como visto anteriormente, o método de controle sintético usa um processo baseado nos dados para escolher os pesos que serão utilizados para construir o contrafactual da unidade tratada. Não há, contudo, uma recomendação específica sobre quais variáveis devem ser utilizadas para estimar esses pesos.

Abadie e Gardeazabal (2003) e Abadie et al. (2010) realizam o pareamento entre a unidade tratada e o seu respectivo controle sintético usando as características observadas que afetam a variável de interesse nos anos antecedentes ao da intervenção. Os pesos dados a cada unidade de controle do *donor pool* são distribuídos entre elas de forma a mimetizar o melhor possível o perfil da unidade tratada antes do tratamento em termos daquelas variáveis observadas. Abadie et al. (2015) e Kleven, Landais, e Saez (2013) usam a média de todas as variáveis de resultado pré-intervenção e outras covariadas. Por sua vez, Billmeier e Nannicini (2013), Bohn, Lofstrom, e Raphael

(2014), Gobillon e Magnac (2016) e Hindrichs (2012) usam somente os valores da variável de resultado para todo o período pré-tratamento. Outros usam somente os valores de resultado para alguns períodos pré-tratamento: Smith (2015) escolhe 4 de 10 períodos, Abadie et al. (2010) escolhem 3 de 19 períodos e Montalvo (2011) escolhe somente os dois períodos antes do tratamento.

Botosaru e Ferman (2019), por outro lado, mostram que o pareamento perfeito das covariadas não constitui um requisito indispensável para a validade do estimador de controle sintético, desde que ocorra um pareamento perfeito de uma série longa da variável de interesse.¹⁷

Ferman, Pinto, e Possebom (2018) mostram que o método de controle sintético é suscetível à busca de especificações, ou seja, o pesquisador pode selecionar especificações que mostrem impactos significativos do ponto de vista estatístico ainda que não haja qualquer efeito. Eles argumentam que não se deve usar a média das variáveis de resultado pré-tratamento ou um número limitado de resultados pré-tratamento, pois não se está desta forma explorando a dinâmica das séries temporais, que é o objetivo último do método de controle sintético. A indicação dada por eles é que sejam usadas especificações com mais defasagens dos resultados pré-tratamento como previsores, ainda que não seja possível determinar em que circunstâncias é melhor usar todas as defasagens ou somente um sub-conjunto dessas defasagens.

Diante disso, os valores do saldo estrutural para cada ano do período pré-tratamento foram usados para calcular os pesos e construir o contrafactual, o saldo estrutural do grupo de controle. Desta forma, o número de períodos utilizados varia de acordo com o ano de adoção do conselho. Por exemplo, a Alemanha teve seu conselho instituído em 2009, iniciando suas atividades no ano seguinte. A Alemanha sintética é então construída de forma a mimetizar os valores exibidos pelo resultado estrutural daquele país para os anos de 2000 a 2009. A partir de 2010, os saldos estruturais efetivamente observados para a Alemanha são subtraídos dos saldos estruturais projetados para a Alemanha sintética para fins de estimação dos efeitos do conselho fiscal.

As tabelas 4 e 5 apresentam as estimativas dos efeitos da adoção do conselho fiscal em cada um dos países com conselhos fiscais fortes e fracos, respectivamente. Encontra-se evidência de que a adoção de conselhos fiscais independentes, quer fortes, quer fracos, não tem efeitos significativos no desempenho fiscal dos países (à exceção de algumas estimativas pontuais, e mesmo assim ao nível de significância de 10%).

¹⁷Os autores estabelecem formalmente dois resultados relevantes em que se apoia essa conclusão:

i) fornecem condições sob as quais o estimador de controle sintético permanece não viesado mesmo quando a hipótese de pareamento perfeito das covariadas é relaxada e substituída pela de pareamento perfeito da variável de interesse; ii) também são fornecidas condições em que o pareamento perfeito da variável de interesse pré-tratamento proporciona um pareamento aproximado das covariadas.

Tabela 4. Efeito estimado por ano de tratamento (Conselhos Fiscais Fortes)

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alemanha	–0,65 (0,25)	–0,31 (0,56)	0,53 (0,39)	0,92 (0,25)	1,61 (0,11)	1,24 (0,22)	1,98 (0,14)	1,95 (0,14)
Austrália	–1,26 (0,25)	–1,44 (0,19)	–1,43 (0,39)	–2,38 (0,17)	–1,49 (0,42)			
Áustria	0,13 (0,92)	2,03 (0,17)	2,45 (0,19)	1,24 (0,44)	1,20 (0,44)			
Chipre	3,67* (0,06)	4,98 (0,17)	3,17 (0,17)	3,74 (0,14)				
Estônia	1,93 (0,25)	6,18 (0,17)	5,07 (0,14)	4,80 (0,14)				
Finlândia	–1,40 (0,69)	–1,55 (0,39)	–1,08 (0,89)	–1,09 (0,86)	–1,78 (0,75)			
França	0,56 (0,44)	1,34 (0,17)	1,66 (0,25)	1,83 (0,17)	1,94 (0,14)			
Holanda	0,56 (0,33)	–0,01 (1,00)	1,03 (0,50)	1,33 (0,31)				
Itália	0,66 (0,22)	0,25 (0,81)	–0,56 (0,67)	–1,00 (0,31)				
Portugal	1,56 (0,50)	2,04 (0,50)	3,45 (0,31)	2,70 (0,53)	2,54 (0,47)	3,04 (0,47)		
Reino Unido	–2,89 (0,47)	–1,51 (0,75)	–3,77 (0,50)	–1,94 (0,78)	–2,46 (0,58)	–2,18 (0,81)	–2,57 (0,78)	–1,37 (0,86)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela 5. Efeito estimado por ano de tratamento (Conselhos Fiscais Fracos)

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
África do Sul	-0,75 (0,14)	-0,76 (0,36)	-0,48 (0,56)	-0,98 (0,17)			
Chile	0,66 (0,14)	1,89 (0,06)	-0,29 (0,69)	-2,10* (0,08)			
Colômbia	1,63* (0,06)	1,09 (0,19)	0,31 (0,53)	1,30 (0,22)	2,45 (0,11)	2,76 (0,11)	
Eslováquia	1,10 (0,50)	1,56 (0,56)	3,19 (0,28)	1,93 (0,50)	-0,52 (0,86)	-0,32 (0,94)	0,92 (0,89)
Espanha	0,96 (0,53)	0,40 (0,89)	-1,69 (0,75)	-0,89 (0,83)			
Irlanda	-2,94 (0,53)	-2,58 (0,64)	-1,07 (0,81)	0,52 (1,00)	-1,07 (0,92)	-0,13 (1,00)	
Luxemburgo	-0,16 (0,89)	-0,11 (0,94)	0,37 (0,94)	0,03 (1,00)			
Malta	-0,25 (0,94)	2,80 (0,19)	5,94* (0,06)				
Peru	-1,69 (0,14)	-1,31 (0,14)	-1,84 (0,11)				

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Como há mais de uma unidade de tratamento (diversos países adotam conselhos fiscais) e o tratamento (adoção dos conselhos) ocorre em diferentes períodos do tempo usa-se também a extensão do estimador de controles sintéticos para múltiplos tratados com heterogeneidade no tempo de intervenção desenvolvida por Cavallo, Galiani, Noy, e Pantano (2013). A Tabela 6 apresenta os efeitos médios estimados a partir dos efeitos para a adoção de cada conselho.¹⁸ Analisados conjuntamente, os

¹⁸Sejam $g = 1, \dots, G$ os países que adotaram CFs e $j = 1, \dots, J$ as unidades de controle. Seja ainda $\hat{\delta}_{gt}$ o efeito da intervenção do CF para o país g no ano t , $t = T_0 \dots T$, computado como $\hat{\delta}_{gt} = \hat{s}_{gt} / \text{RMSPE}_g$. A divisão dos efeitos estimados pela qualidade do ajuste (RMSPE) tem a finalidade de dar mais peso aos resultados obtidos de pareamentos mais próximos. Suponhamos agora que queiramos inferir o efeito para o primeiro ano de tratamento (para simplificar a exposição, omitimos o subscrito t , pois os anos de adoção do CF podem variar entre os países). A média para esse efeito entre as unidades tratadas é $\bar{\delta} = G^{-1} \sum_{g=1}^G \hat{\delta}_g$. Chamemos de $\hat{\delta}_g^{\text{PL}}$ o conjunto dos efeitos de placebo estimados para o país g , supondo que cada unidade de controle receba o CF no mesmo ano que g . É então necessário comparar $\bar{\delta}$ com a distribuição das médias dos efeitos de diferentes grupos de placebo de tamanho G , $\bar{\delta}^{\text{PL}}$. Foram formados aleatoriamente e sem reposição 1000 grupos de placebo para o teste, extraídos da totalidade de países que não adotaram CFs. Nessa inferência, o

Tabela 6. Efeito médio estimado por ano de tratamento (Cavallo et al., 2013)

Var. Dep.: SB	Efeito médio por ano de tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conselhos Fortes	0,19 (0,73)	0,73 (0,21)	0,88 (0,18)	0,78 (0,33)	0,86 (0,39)	1,22 (0,35)	1,49 (0,30)	1,59 (0,34)
Conselhos Fracos	0,13 (0,75)	0,54 (0,43)	0,09 (0,90)	-0,38 (0,70)	1,54 (0,34)	1,87 (0,27)	0,92 (0,62)	
Todos os Conselhos	0,16 (0,47)	0,63* (0,07)	0,50 (0,36)	0,28 (0,77)	1,06 (0,22)	1,53 (0,11)	1,38 (0,33)	1,59 (0,31)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) efeito médio estimado a partir da combinação aleatória de 1000 outros grupos de igual tamanho ao grupo tratado extraídos dos países de controle, sem reposição (Cavallo et al., 2013); iv) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

efeitos dos conselhos fiscais não mostram significância para qualquer dos grupos selecionados.

Conclui-se assim que a informação de estudos de caso obtida com a metodologia de controle sintético traz à tona evidência importante. Na verdade, estimativas dos efeitos causais médios devem ser vistos com extrema cautela. O método de controle sintético ao gerar estimativas de um caso (país) de cada vez permite explorar a heterogeneidade dos efeitos da adoção de conselhos fiscais, ao contrário do que ocorre com a estimação de diferença-em-diferenças, em que em geral são apresentados somente os efeitos médios. No presente caso, a estimação de efeitos individuais é particularmente importante uma vez que o desenho dos conselhos fiscais é muito diferente, assim como o ambiente em que a adoção dos conselhos ocorreu, o que certamente afeta seus impactos.

Pela sua natureza, a variável de saldo primário estrutural não capta oscilações de curto prazo/cíclicas. Assim, não é possível descartar que os conselhos fiscais não impactem o resultado estrutural, mas que tenham algum impacto sobre o seu componente transitório. A fim de verificar se esse é o caso, as estimativas foram refeitas usando-se como variável dependente o saldo primário propriamente dito. Os resultados continuam apontando a ausência de influência dos conselhos sobre o desempenho fiscal (ver seção A.2).

p-valor é computado como:

$$p - \text{valor} = \Pr(|\bar{\delta}^{\text{PL}}| \geq |\bar{\delta}_g|) = \frac{\sum_{i=1}^{1000} \mathbb{1}(|\bar{\delta}^{\text{PL}(i)}| \geq |\bar{\delta}_g|)}{1000},$$

onde $\mathbb{1}(\cdot)$ é a função indicadora para o evento \cdot e i indica o efeito-placebo para cada um dos 1.000 grupos de placebo. Rejeita-se a hipótese nula de não efeito se p for menor que algum nível de significância previamente especificado.

Para verificar quão bem o país sintético mimetiza o país tratado antes do tratamento, ou seja, avaliar se o país de controle criado pelo método de controle sintético é um contrafactual adequado, segue-se [Abadie et al. \(2010\)](#) e calcula-se o RMSPE. Os resultados são apresentados na [Tabela 7](#). O ajuste mostra boa qualidade, com mediana de 0,73 e pequena variabilidade — desvio-padrão de 0,65.

Tabela 7. Qualidade do ajuste pré-tratamento do controle sintético

	RMSPE		RMSPE
Alemanha	0,24	Eslováquia	0,90
Chile	0,25	Chipre	0,95
Colômbia	0,27	Malta	1,08
África do Sul	0,29	Portugal	1,10
Itália	0,40	Estônia	1,28
Peru	0,41	Holanda	1,52
França	0,45	Itália	1,54
Holanda	0,51	Luxemburgo	1,77
Austrália	0,55	Malta	1,93
Áustria	0,55	Peru	2,47

5. Testes adicionais

São conduzidos cinco testes adicionais para verificar se de fato a adoção de conselhos fiscais independentes não é eficaz: i) teste da razão pós/pré RMSPE; ii) restrição do *donor pool*; iii) uso da variação do saldo estrutural como variável de resultado; iv) substituição dos regressores por covariadas fiscais; v) postergação da adoção dos conselhos em dois anos (efeito maturidade).

Primeiro, aplica-se o teste da razão pós/pré RMSPE ([Abadie et al., 2010](#)) para cada um dos países avaliados. A vantagem desse teste consiste na possibilidade do exame conjunto de todos os anos de atuação do conselho, para cada país. Uma razão de grande magnitude do tratado em relação aos placebos constitui evidência de efetividade da instituição dos CFs, o que pode ocorrer mesmo na ausência de significância do efeito anual. Os resultados confirmam os obtidos anteriormente, ou seja, para nenhum dos países o coeficiente da adoção de conselhos fiscais apresentou efeito significativo, como pode ser visto na [Tabela 8](#).

Inicialmente foram incluídos no *donor pool* todos os países com dados disponíveis ao redor do mundo. No intuito de gerar maior homogeneidade entre os países tratados e controles, principalmente com relação ao nível de desenvolvimento e qualidade das instituições, o grupo de controle foi restringido aos países com renda per capita similar à dos países tratados ([Tapsoba, 2012](#)). Assim, passaram a fazer

Tabela 8. Qualidade do ajuste pré-tratamento do controle sintético

	Pré_RMSPE	Pós_RMSPE	Razão Pré/Pós	p-valor
CFs Fortes				
Alemanha	0,24	1,39	5,90	0,24
Austrália	0,55	1,84	3,34	0,27
Áustria	0,55	1,81	3,32	0,27
Chipre	0,95	4,56	4,79	0,14
Estônia	1,28	5,50	4,29	0,14
Finlândia	1,52	1,57	1,04	0,95
França	0,45	1,73	3,87	0,19
Holanda	0,51	1,02	2,02	0,54
Itália	0,40	0,77	1,96	0,54
Portugal	1,10	2,88	2,62	0,51
Reino Unido	1,93	2,62	1,35	0,95
CFs Fracos				
África do Sul	0,29	0,88	3,04	0,27
Chile	0,25	1,68	6,78	0,14
Colômbia	0,27	1,97	7,16	0,14
Eslováquia	0,90	1,77	1,96	0,65
Espanha	1,54	1,25	0,81	0,92
Irlanda	2,47	1,89	0,77	1,00
Luxemburgo	1,77	0,24	0,14	1,00
Malta	1,08	4,65	4,31	0,22
Portugal	0,41	2,00	4,83	0,19

parte do *donor pool* somente os países com renda per capita maior do que a Colômbia, o país tratado com a menor renda per capita.¹⁹ Os resultados são apresentados na seção A.3 e mostram efeitos não significativos para todos países, tanto na avaliação individual como conjunta (Cavallo et al., 2013), corroborando os achados da análise principal.

O terceiro teste de robustez substitui o resultado estrutural em nível, utilizado na estimação principal, pela sua variação ano a ano. Neste caso, a hipótese é que a melhora na qualidade do ajuste fiscal após a adoção do conselho se manifestou na variação do resultado fiscal (melhora dos superávits ou redução dos déficits) e não no seu nível. Os resultados obtidos encontram-se na seção A.4 e, mais uma vez, mostram que, excluídos alguns poucos anos específicos, permanece a ausência de significância para os efeitos encontrados.

¹⁹O donor pool passa a ser composto por: Argentina, Brasil, Bulgária, Cazaquistão, Croácia, República Checa, República Dominicana, Granada, Hong Kong, Hungria, Islândia, Israel, Japão, Malásia, Nova Zelândia, Noruega, Polônia, Rússia, Cingapura, Eslovênia, Suriname, Suíça, Tailândia, Turquia e Uruguai.

Como visto anteriormente, os regressores consistem nas variáveis defasadas do saldo estrutural, seguindo pesquisas recentes sobre a melhor qualidade dessa especificação em trajetórias razoavelmente longas de pré-tratamento (Botosaru & Ferman, 2019). Contudo, para reforçar os resultados encontrados, o contrafactual foi construído utilizando-se sete covariadas comumente empregadas na literatura como capazes de influir na trajetória do resultado fiscal: a dívida pública (defasada um ano), o nível e a variação do PIB per capita, a taxa de inflação, o grau de abertura comercial, a razão de dependentes na população e o número de anos de eleição.²⁰ Na construção do controle sintético, empregamos a média dos cinco anos pré-tratamento. Vale ressaltar a expressiva redução na qualidade do ajuste pré-tratamento (medido pelo RMSPE) quando comparado com o obtido anteriormente, bem como o aumento na magnitude dos efeitos encontrados. Ainda assim, os efeitos mantêm-se sem significância para todos os países, com algumas exceções pontuais (ver seção A.5).

Em nosso quinto e último teste de robustez a adoção do conselho fiscal é postergada em dois anos. Isto porque é natural imaginar que seja necessária a decorrência de algum tempo para que efetivamente surja algum efeito restritivo do conselho fiscal sobre o comportamento da política fiscal. No intuito de melhor capturar esse “efeito maturidade”, realiza-se uma segunda rodada de estimações assumindo que o primeiro ano de atuação do conselho ocorreu somente dois anos após sua implementação efetiva. Uma vantagem adicional deste teste consiste em alongar o período de pré-tratamento, conferindo um ajuste mais consistente ao modelo (Botosaru & Ferman, 2019). Ainda assim, os coeficientes mantiveram sua não significância estatística, como pode ser visto na seção A.6.

6. Conclusões

O objetivo deste artigo é mensurar o impacto de reformas institucionais no setor público, mais precisamente da adoção de conselhos fiscais sobre o resultado primário estrutural. Utilizando um painel de 56 países durante o período 2000 a 2017, utiliza-se o método de diferença-em-diferenças tradicional e uma versão modificada do método de controle sintético para múltiplos tratados com heterogeneidade no ano de adoção da instituição fiscal independente.

O grupo de tratamento é composto por 20 países, que se diferenciam de acordo com a qualidade dos seus conselhos fiscais, classificada em forte e fraca. No primeiro grupo, composto por 11 países, incluem-se apenas os conselhos que elaboram ou realizam previsões macroeconômicas, monitoram o cumprimento de regras fiscais, têm impacto na mídia, participam do processo orçamentário e têm independência. O segundo grupo é formado por 9 países, nos quais uma ou mais dessas características

²⁰ As fontes e descrição dessas variáveis podem ser encontradas na Tabela A-1.

estão ausentes. O comportamento fiscal dos governos é medido pelo resultado estrutural como proporção do PIB potencial.

Os resultados encontrados apontam para a ausência de qualquer influência dos conselhos fiscais sobre a trajetória do resultado estrutural, tanto ao nível individual dos países como nas avaliações de grupo.

Tais achados revelam uma contribuição mais modesta ou inexistente dos conselhos para o desempenho fiscal dos países, na comparação com a encontrada em trabalhos anteriores. Atribuímos esse resultado à melhor qualidade no tratamento da endogeneidade efetuado pelo método de controle sintético.

Em termos de implicações de política, os resultados sugerem que a implementação de instituições fiscais independentes *per se* não é capaz de garantir a condução adequada da política fiscal, mesmo no médio prazo, quando os efeitos da reputação adquirida poderiam gerar efeitos mais pronunciados (“efeito maturidade”). Propõe-se dessa forma uma mudança de enfoque nas pesquisas sobre a efetividade dessas instituições, que leve em conta as particularidades de seus objetivos e canais de atuação, diferenciando-as assim de suas antecessoras, as regras fiscais. Enquanto estas atuam e são avaliadas por seus impactos diretos sobre o resultado fiscal, as formulações mais recentes de desenho ótimo dos conselhos (Diretiva de dezembro de 2017 da União Europeia) preconizam que estes monitorem e avaliem publicamente: (i) a adequação e observância das regras fiscais nacionais; (ii) a qualidade das previsões oficiais; (iii) a ocorrência ou a cessação de circunstâncias que desencadeiem cláusulas de escape sob as regras. A estas recomendações também cabe adicionar a capacidade dos conselhos em dar ampla publicidade a suas avaliações, gerando impacto sobre a mídia. Tais devem ser os objetivos a serem avaliados numa atuação apropriada das instituições fiscais independentes.

Referências bibliográficas

- Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490), 493–505. <http://dx.doi.org/10.1198/jasa.2009.ap08746>
- Abadie, A., Diamond, A., & Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2), 495–510. <http://dx.doi.org/10.1111/ajps.12116>
- Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the Basque Country. *American Economic Review*, 93(1), 113–132. <http://dx.doi.org/10.1257/000282803321455188>
- Badinger, H., & Reuter, W. H. (2017). The case for fiscal rules. *Economic Modeling*, 60, 334–343. <http://dx.doi.org/10.1016/j.econmod.2016.09.028>

- Beetsma, R., Debrun, X., Fang, X., Kim, Y., Lledó, V., Mbaye, S., & Zhang, X. (2018, março). *Independent fiscal councils: Recent trends and performance* (Working Paper N° WP/18/68). Washington, DC: IMF. <https://www.imf.org/en/Publications/WP/Issues/2018/03/23/Independent-Fiscal-Councils-Recent-Trends-and-Performance-45726>
- Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *The Quarterly Journal of Economics*, 119(1), 249–275. <http://dx.doi.org/10.1162/003355304772839588>
- Billmeier, A., & Nannicini, T. (2013). Assessing economic liberalization episodes: A synthetic control approach. *The Review of Economics and Statistics*, 95(3), 983–1001. http://dx.doi.org/10.1162/REST_a_00324
- Bohn, S., Lofstrom, M., & Raphael, S. (2014). Did the 2007 Legal Arizona Workers Act reduce the state's unauthorized immigrant population? *The Review of Economics and Statistics*, 96(2), 258–269. http://dx.doi.org/10.1162/REST_a_00429
- Botosaru, I., & Ferman, B. (2019). On the role of covariates in the synthetic control method. *The Econometrics Journal*, 22(2), 117–130. <http://dx.doi.org/10.1093/ectj/utz001>
- Calmfors, L., & Wren-Lewis, C. (2011, fevereiro). *The role of independent fiscal policy institutions* (Working Paper N° 3367). Munich: CESifo. <https://www.cesifo.org/node/16961>
- Cavallo, E., Galiani, S., Noy, I., & Pantano, J. (2013). Catastrophic natural disasters and economic growth. *The Review of Economic and Statistics*, 95(5), 1549–1561. http://dx.doi.org/10.1162/REST_a_00413
- Coletta, G., & Infantino, G. (2015). *Do fiscal councils impact fiscal performance?* Department of Treasure of Italy.
- Debrun, X., Gérard, M., & Harris, J. (2011). Fiscal policies in crisis mode: Has the time for fiscal councils come at last? *In DG ECFIN Workshop on Fiscal Policies during the Crisis*.
- Debrun, X., Hauner, D., & Kumar, M. S. (2009). Independent fiscal agencies. *Journal of Economic Surveys*, 23(1), 44–81. <http://dx.doi.org/10.1111/j.1467-6419.2008.00556.x>
- Debrun, X., & Kinda, T. (2017). Strengthening post-crisis fiscal credibility: Fiscal councils on the rise – a new dataset. *Fiscal Studies*, 38(4), 667–700. <http://dx.doi.org/10.1111/1475-5890.12130>
- Debrun, X., & Kumar, M. S. (2007, julho). *The discipline-enhancing role of fiscal institutions: Theory and empirical evidence* (Working Paper N° WP/07/171). Washington, DC: IMF. <https://www.imf.org/external/pubs/ft/wp/2007/wp07171.pdf>
- Debrun, X., & Kumar, M. S. (2008, junho). Fiscal rules, fiscal councils and all that: Commitment devices, signaling tools or smokescreens? *In Fiscal policy: Current issues and challenges — proceedings of the 9th Banca d'Italia Workshop on Public Finance*. Rome: Banca d'Italia. https://www.bancaditalia.it/pubblicazioni/altri-atti-convegni/2007-fiscal-policy/Debrun_Kumar.pdf?language_id=1
- Eichengreen, B., Hausmann, R., & Von Hagen, J. (1999). Reforming budgetary institutions in Latin America: The case for a national fiscal council. *Open Economies Review*, 10, 415–442. <http://dx.doi.org/10.1023/A:1008337818753>

- Ferman, B., Pinto, C., & Possebom, V.** (2018, março). Cherry picking with synthetic controls. *MPRA*, 85138. <https://mpra.ub.uni-muenchen.de/id/eprint/85138>
- Frankel, J., & Schreger, J.** (2012). Over-optimistic official forecasts and fiscal rules in the Eurozone. *Review of World Economics*, 149, 247–272. <http://dx.doi.org/10.1007/s10290-013-0150-9>
- Gobillon, L., & Magnac, T.** (2016). Regional policy evaluation: Interactive fixed effects and synthetic controls. *Review of Economics and Statistics*, 98(3), 535–551. http://dx.doi.org/10.1162/REST_a_00537
- Hallerberg, M., Strauch, R., & Von Hagen, J.** (2009). *Fiscal governance in Europe*. Cambridge: Cambridge University Press.
- Heinemann, F., Osterloh, S., & Kalb, A.** (2014). Sovereign risk premia: The link between fiscal rules and stability culture. *Journal of International Money and Finance*, 41, 110–127. <http://dx.doi.org/10.1016/j.jimonfin.2013.11.002>
- Hindrichs, P.** (2012). The effects of affirmative action bans on college enrollment, educational attainment, and the demographic composition of universities. *Review of Economics and Statistics*, 94(3), 712–722. http://dx.doi.org/10.1162/REST_a_00170
- Jonung, L., & Larch, M.** (2006). Improving fiscal policy in the EU: The case for independent forecasts. *Economic Policy*, 21. <https://www.jstor.org/stable/3874052>
- Kleven, H. J., Landais, C., & Saez, E.** (2013). Taxation and international migration of superstars: Evidence from European football market. *American Economic Review*, 103(5), 1892–1924. <http://dx.doi.org/10.1257/aer.103.5.1892>
- Montalvo, J. G.** (2011). Voting after the bombings: A natural experiment and the effect of terrorist attacks on democratic elections. *Review of Economic and Statistics*, 93(4), 1146–1154. http://dx.doi.org/10.1162/REST_a_00115
- Nerlich, C., & Reuter, W. H.** (2013, setembro). *The design of national frameworks and their budgetary impacts* (Working Paper N° 1588). Frankfurt am Main, Germany: European Central Bank. <https://www.ecb.europa.eu/pub/pdf/scpwpwps/ecbwp1588.pdf>
- Poterba, J.** (1996, abril). *Do budget rules work?* (Working Paper N° 5550). NBER. <http://dx.doi.org/10.3386/w5550>
- Reuter, W. H.** (2015). National numerical fiscal rules: Not complied with, but still effective? *European Journal of Political Economy*, 39, 67–81. <http://dx.doi.org/10.1016/j.ejpoleco.2015.04.002>
- Smith, B.** (2015). The resource curse exorcised: Evidence from a panel of countries. *Journal of Developing Economics*, 116, 57–73. <http://dx.doi.org/10.1016/j.jdeveco.2015.04.001>
- Tapsoba, R.** (2012). Do national numerical fiscal rules really shape fiscal behaviour in developing countries? A treatment effect evaluation. *Economic Modelling*, 29(4), 1356–1369. <http://dx.doi.org/10.1016/j.econmod.2012.03.003>
- Von Hagen, J.** (1992). *Budgeting procedures and fiscal performance in the European communities* (Working Paper N° 96). European Commission.
- Von Hagen, J., & Harden, I.** (1995). Budget processes and commitment to fiscal discipline. *European Economic Review*, 39(3-4), 771–779. [http://dx.doi.org/10.1016/0014-2921\(94\)00084-D](http://dx.doi.org/10.1016/0014-2921(94)00084-D)

Apêndice A.

A.1 Fontes e descrição dos dados

Tabela A-1. Fontes e descrição dos dados

Variável	Descrição	Fonte
Conselho Fiscal	Dummy assumindo valor igual a 1 se em determinado ano o país possui Conselho Fiscal	Fiscal Council Dataset – IMF December, 2016
SB	Saldo estrutural do governo geral Percentual do PIB potencial	
Dívida Pública	Dívida bruta do governo geral Percentual do PIB	
Variação do PIB per capita	Variação do PIB per capita a preços constantes Paridade do poder de compra; em dólares de 2011	WEO Database – IMF October, 2017
Inflação	Média da inflação de preços de consumo Variação percentual	
Taxa de Desemprego	Taxa de desemprego Percentual da população ativa	
Ano de Eleição	Dummy assumindo valor igual a 1 se em determinado ano houve eleição para o executivo e/ou legislativo	
Ideologia do Executivo	Orientação do partido no executino com relação à política econômica: Direita (1), Esquerda (3), Centro (2)	Database of Political Institutions DPI 2017, BID
Fragmentação do Congresso	Probabilidade de dois deputados tomados aleatoriamente no Congresso serem de partidos diferentes da base do governo	
Estabilidade Política	Índice variando de 0 a 1 e crescente com a capacidade do governo em se manter no cargo e realizar os programas anunciados	
Qualidade da Burocracia	Índice variando de 0 a 1 e crescente com a força institucional e expertise da burocracia em governar sem causar mudanças drásticas	International Country Risk Guide ICRG 2017 World Bank version
Controle da Corrupção	Índice variando de 0 a 1 e crescente com a capacidade das instituições em coibir a prática da corrupção	
Abertura Comercial	Soma das exportações e importações Percentual do PIB	World Bank Database National Accounts Data
Razão de Dependência	População com menos de 15 e mais de 64 anos Percentual da população com idade entre 15 e 64 anos	
Regra Fiscal	Dummy assumindo valor igual a 1 se em determinado ano o país possui uma regra fiscal numérica nacional em vigor	Fiscal Rules Database IMF December, 2016

A.2 Efeito estimado por ano de tratamento com a utilização do saldo primário efetivo como variável dependente

Tabela A-2. Conselhos fiscais fortes – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alemanha	-1,62 (0,26)	1,87 (0,34)	2,34 (0,16)	1,25 (0,37)	1,61 (0,24)	1,63 (0,29)	1,86 (0,32)	1,98 (0,26)
Austrália	0,06 (0,97)	-0,52 (0,32)	-2,05 (0,13)	-1,82 (0,13)	-0,62 (0,50)			
Áustria	2,59 (0,13)	0,11 (0,89)	-0,70 (0,76)	-0,70 (0,71)	0,07 (1,00)			
Chipre	4,59* (0,05)	2,64 (0,32)	4,43 (0,16)	5,02 (0,08)				
Estônia	0,47 (0,47)	0,07 (0,92)	-1,40 (0,34)	0,37 (0,76)				
Finlândia	-0,70 (0,84)	-0,75 (0,66)	-0,21 (0,95)	2,11 (0,63)	0,81 (0,87)			
França	1,32 (0,13)	0,71 (0,26)	-0,95 (0,39)	-0,21 (0,84)	0,22 (0,87)			
Holanda	0,62 (0,53)	0,63 (0,71)	3,18 (0,16)	3,31 (0,08)				
Itália	0,71 (0,32)	-0,32 (0,79)	1,07 (0,55)	1,87 (0,13)				
Portugal	2,31 (0,37)	2,90 (0,42)	-0,28 (0,89)	-0,08 (1,00)	2,82 (0,53)	3,55 (0,29)		
Reino Unido	-2,81 (0,39)	-2,19 (0,58)	-2,16 (0,55)	-1,68 (0,71)	-1,36 (0,71)	-2,44 (0,63)	-1,14 (0,87)	0,99 (0,79)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-3. Conselhos fiscais fracos – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
África do Sul	1,04 (0,16)	0,09 (0,89)	1,59 (0,16)	0,73 (0,42)				
Chile	-0,11 (0,97)	0,19 (0,92)	-0,79 (0,79)	-1,89 (0,61)				
Colômbia	1,94** (0,03)	2,20** (0,03)	2,03*** (0,00)	0,71 (0,18)	0,52 (0,37)	0,62 (0,18)		
Eslováquia	2,44 (0,53)	2,75 (0,32)	5,18 (0,18)	4,01 (0,21)	2,51 (0,53)	2,69 (0,61)	2,72 (0,45)	
Espanha	-1,99 (0,55)	-2,15 (0,71)	-3,62 (0,61)	-0,12 (1,00)				
Irlanda	-2,36 (0,95)	-1,06 (0,97)	-0,83 (0,95)	-0,47 (1,00)	-5,45 (0,82)	0,75 (1,00)		
Luxemburgo	1,98 (0,21)	0,43 (0,84)	2,18 (0,58)	1,08 (0,66)				
Malta	-0,02 (0,97)	2,67 (0,34)	2,18 (0,29)					
Peru	-1,08 (0,47)	-0,90 (0,61)	-1,37 (0,39)					

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-4. Efeito médio por ano de tratamento (Cavallo et al., 2013)

Var. Dep.: SB	Efeito médio por ano de tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conselhos Fortes	0,58 (0,56)	0,44 (0,66)	0,11 (0,93)	0,58 (0,62)	0,38 (0,70)	1,30 (0,32)	1,22 (0,49)	1,77*** 0,00
Conselhos Fracos	1,12*** (0,00)	1,21** (0,01)	1,46 (0,15)	0,74 (0,24)	0,61 (0,35)	0,82 (0,20)	2,72 (0,27)	
Todos os Conselhos	0,82 (0,74)	0,78 (0,76)	0,71 (0,69)	0,64 (0,78)	0,47 (0,80)	0,98 (0,48)	1,50 (0,50)	1,77*** 0,00

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) efeito médio estimado a partir da combinação aleatória de 1000 outros grupos de igual tamanho ao grupo tratado extraídos dos países de controle, sem reposição (Cavallo et al., 2013); iv) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

A.3 Testes de robustez com restrição do *donor pool*

Tabela A-5. Conselhos fiscais fortes – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alemanha	–0,89 (0,24)	–0,17 (0,88)	0,89 (0,28)	1,67 (0,12)	2,31 (0,08)	1,24 (0,28)	1,67 (0,12)	1,97 (0,12)
Austrália	–1,26 (0,20)	–1,44 (0,20)	–1,43 (0,32)	–2,38 (0,16)	–1,49 (0,40)			
Áustria	0,24 (0,88)	1,94 (0,16)	2,84 (0,16)	1,41 (0,36)	1,35 (0,44)			
Chipre	3,67** (0,04)	4,98 (0,12)	3,17 (0,20)	3,74 (0,16)				
Estônia	1,38 (0,40)	6,04 (0,12)	4,89 (0,20)	5,07 (0,16)				
Finlândia	–1,57 (0,60)	–1,10 (0,72)	0,04 (1,00)	–0,40 (1,00)	–1,00 (0,88)			
França	0,41 (0,72)	0,54 (0,48)	0,33 (0,68)	–0,05 (1,00)	0,12 (1,00)			
Holanda	0,67 (0,32)	0,20 (0,80)	1,24 (0,32)	1,50 (0,32)				
Itália	0,82 (0,32)	1,25 (0,28)	0,26 (0,88)	0,28 (0,96)				
Portugal	2,05 (0,36)	2,84 (0,24)	3,27 (0,28)	1,81 (0,52)	1,98 (0,60)	2,74 (0,44)		
Reino Unido	–2,89 (0,56)	–1,51 (0,64)	–3,77 (0,40)	–1,94 (0,88)	–2,46 (0,60)	–2,18 (0,60)	–2,57 (0,76)	–1,37 (0,88)

Notas: i) 25 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-6. Conselhos fiscais fracos – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
África do Sul	–0,26 (0,64)	0,20 (0,80)	0,23 (0,84)	0,09 (1,00)			
Chile	0,45 (0,32)	2,17** (0,04)	0,04 (0,96)	–1,26 (0,16)			
Colômbia	1,35 (0,12)	0,92 (0,12)	0,08 (0,96)	1,00 (0,20)	2,56 (0,08)	2,79 (0,08)	
Espanha	0,57 (0,80)	–0,12 (0,92)	–2,22 (0,76)	–1,24 (0,88)			
Eslováquia	0,86 (0,68)	1,65 (0,60)	3,51 (0,16)	1,46 (0,64)	0,31 (0,88)	0,08 (0,96)	1,65 (0,64)
Irlanda	–3,07 (0,48)	–2,66 (0,64)	–1,18 (0,88)	0,31 (0,92)	–1,35 (0,84)	–0,43 (1,00)	
Luxemburgo	0,00 (1,00)	0,45 (0,84)	0,01 (1,00)	–0,01 (1,00)			
Malta	–0,13 (0,84)	2,49 (0,44)	6,05 (0,12)				
Peru	–1,73 (0,12)	–2,02 (0,12)	–2,59* (0,08)				

Notas: i) 25 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-7. Efeito médio por ano de tratamento (Cavallo et al., 2013)

Var. Dep.: SB	Efeito médio por ano de tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conselhos Fortes	0,18 (0,84)	0,89 (0,18)	1,01 (0,27)	0,91 (0,35)	0,65 (0,68)	1,10 (0,57)	1,09 (0,59)	1,51 (0,44)
Conselhos Fracos	0,11 (0,80)	0,65 (0,31)	0,15 (0,82)	0,00 (0,99)	1,79 (0,30)	2,00 (0,31)	1,65 (0,38)	
Todos os Conselhos	0,15 (0,87)	0,77*** (0,01)	0,60 (0,45)	0,52 (0,67)	1,00 (0,64)	1,56 (0,39)	1,19 (0,65)	1,51 (0,50)

Notas: i) 25 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) efeito médio estimado a partir da combinação aleatória de 1000 outros grupos de igual tamanho ao grupo tratado extraídos dos países de controle, sem reposição (Cavallo et al., 2013); iv) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

A.4 Teste de robustez com variação do resultado estrutural

Tabela A-8. Conselhos fiscais fortes – Efeito estimado por ano de pós-tratamento

Var. Dep.: Variação SB	Efeito estimado por ano de pós-tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alemanha	–0,97 (0,19)	1,69 (0,11)	0,94 (0,25)	0,13 (0,64)	0,29 (0,53)	–0,69 (0,36)	0,58 (0,31)	0,07 (0,86)
Austrália	0,86 (0,11)	0,46 (0,39)	0,35 (0,61)	–1,55 *** (0,00)	0,74 (0,11)			
Áustria	0,30 (0,75)	1,39 (0,28)	0,30 (0,81)	–1,10 (0,39)	0,02 (1,00)			
Chipre	1,33 (0,42)	1,56 (0,56)	–1,61 (0,56)	0,17 (0,94)				
Estônia	0,81 (0,58)	–1,79 (0,53)	0,24 (0,92)	–1,30 (0,42)				
Finlândia	0,21 (0,81)	0,45 (0,53)	0,27 (0,81)	0,07 (0,94)	–0,11 (0,94)			
França	1,38 ** (0,03)	0,99 (0,11)	0,59 (0,28)	–0,67 (0,25)	0,56 (0,19)			
Holanda	–0,24 (0,83)	–0,36 (0,75)	1,46 (0,31)	–0,15 (0,86)				
Itália	–0,35 (0,50)	–0,52 (0,58)	–1,52 (0,19)	–0,65 (0,36)				
Portugal	4,11 * (0,06)	1,27 (0,19)	0,77 (0,50)	–0,60 (0,67)	0,87 (0,61)	0,23 (0,86)		
Reino Unido	0,28 (0,92)	1,01 (0,61)	–1,37 (0,47)	1,88 (0,19)	–0,32 (0,83)	0,78 (0,67)	0,03 (0,97)	1,21 (0,44)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-9. Conselhos fiscais fracos – Efeito estimado por ano de pós-tratamento

Var. Dep.: Variação SB	Efeito estimado por ano de pós-tratamento						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
África do Sul	-0,82 *	-0,22	0,15	-0,29			
	(0,08)	(0,58)	(0,75)	(0,36)			
Chile	0,26	1,46	-2,01	-1,65 ***			
	(0,44)	(0,14)	(0,03)	(0,00)			
Colômbia	1,20	-0,61	-1,01	0,50	1,21	0,21	
	(0,19)	(0,25)	(0,31)	(0,53)	(0,22)	(0,75)	
Espanha	0,27	-0,81	-2,04	0,79			
	(0,89)	(0,58)	(0,28)	(0,53)			
Eslováquia	2,77	1,06	2,17	-0,58	-0,52	0,30	1,24
	(0,17)	(0,47)	(0,11)	(0,56)	(0,72)	(0,75)	(0,31)
Irlanda	0,01	1,81	2,17	1,62	-2,07	0,34	
	(1,00)	(0,53)	(0,47)	(0,64)	(0,61)	(0,94)	
Luxemburgo	-1,95	-0,77	0,62	0,50			
	(0,25)	(0,64)	(0,75)	(0,72)			
Malta	-0,15	1,91	3,09				
	(0,94)	(0,58)	(0,11)				
Peru	-1,04	-0,41	-0,86				
	(0,28)	(0,69)	(0,11)				

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-10. Efeito médio por ano de tratamento (Cavallo et al., 2013)

Var. Dep.: SB	Efeito médio por ano de tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Conselhos Fortes	0,49	0,75	0,25	-0,43 **	0,38	-0,28	0,47	0,30
	(0,46)	(0,22)	(0,71)	(0,02)	(0,61)	(0,81)	(0,67)	(0,47)
Conselhos Fracos	-0,02	0,20	-0,32	-0,29	0,23	0,26	1,23 ***	
	(0,89)	(0,60)	(0,33)	(0,63)	(0,78)	(0,63)	(0,00)	
Todos os Conselhos	0,26	0,50	-0,01	-0,38	0,34	-0,07	0,63	0,30
	(0,61)	(0,24)	(0,99)	(0,54)	(0,85)	(0,97)	(0,83)	(0,35)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) efeito médio estimado a partir da combinação aleatória de 1000 outros grupos de igual tamanho ao grupo tratado extraídos dos países de controle, sem reposição (Cavallo et al., 2013); iv) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

A.5 Teste de robustez com covariadas nos regressores

Tabela A-11. Conselhos fiscais fortes – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Alemanha	2,42 (0,23)	2,82 (0,17)	3,96 (0,06)	4,22 (0,11)	4,84 (0,11)	4,27 (0,23)	4,72 (0,14)	4,76 (0,11)
Austrália	-0,30 (0,94)	-0,10 (0,94)	0,62 (0,71)	1,04 (0,60)	1,85 (0,49)			
Áustria	1,89 (0,26)	3,12 (0,11)	4,09** (0,03)	3,12 (0,17)	2,86 (0,14)			
Chipre	2,98 (0,26)	4,27 (0,14)	1,44 (0,49)	2,19 (0,49)				
Estônia	2,1 (0,43)	1,73 (0,57)	-0,03 (0,97)	-0,80 (0,86)				
Finlândia	3,44 (0,51)	3,47 (0,40)	3,94 (0,49)	4,34 (0,46)	4,18 (0,51)			
França	0,85 (0,63)	0,59 (0,69)	-0,02 (1,00)	-0,90 (0,66)	-0,75 (0,77)			
Holanda	1,86 (0,37)	1,90 (0,43)	4,01 (0,14)	4,05 (0,14)				
Itália	2,90 (0,26)	1,87 (0,43)	-0,23 (0,94)	-0,52 (0,86)				
Portugal	0,37 (0,86)	1,16 (0,80)	1,90 (0,60)	0,23 (0,94)	-0,44 (0,94)	-0,34 (0,97)		
Reino Unido	-2,49 (0,66)	-1,57 (0,74)	-3,52 (0,49)	-1,92 (0,74)	-2,98 (0,54)	-2,67 (0,54)	-1,84 (0,74)	-0,84 (0,97)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-12. Conselhos fiscais fracos – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
África do Sul	-2,14 (0,20)	-3,09** (0,03)	-3,77 (0,16)	-3,82 (0,19)			
Chile	-1,55 (0,40)	5,76 (0,03)	2,65 (0,37)	2,32 (0,49)			
Colômbia	0,47 (0,71)	-0,19 (0,91)	-0,67 (0,49)	-1,03 (0,43)	0,36 (0,80)	0,05 (0,97)	
Espanha	2,42 (0,43)	0,22 (0,97)	-1,45 (0,74)	-1,43 (0,77)			
Eslováquia	-2,65 (0,71)	-3,70 (0,54)	-1,73 (0,77)	-1,68 (0,71)	-1,99 (0,74)	-2,74 (0,66)	-1,48 (0,89)
Irlanda	-5,01 (0,34)	-3,56 (0,46)	-1,80 (0,66)	-0,44 (0,91)	-0,26 (0,97)	-0,77 (0,94)	
Luxemburgo	1,35 (0,43)	2,91 (0,29)	3,86 (0,17)	3,30 (0,17)			
Malta	-4,10 (0,43)	0,49 (0,94)	3,04 (0,60)				
Peru	1,83 (0,31)	1,41 (0,46)	0,48 (0,80)				

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

A.6 Teste de robustez para efeito maturidade

Tabela A-13. Conselhos fiscais fortes – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento					
	(1)	(2)	(3)	(4)	(5)	(6)
Alemanha	0,71 (0,22)	1,04 (0,17)	1,79** (0,06)	1,48 (0,17)	2,17 (0,11)	2,12 (0,11)
Austrália	-0,03 (0,97)	-0,64 (0,67)	0,07 (0,97)			
Áustria	2,04 (0,17)	1,00 (0,61)	1,20 (0,42)			
Chipre	1,96 (0,53)	2,78 (0,39)				
Estônia	0,59 (0,81)	-0,25 (0,97)				
Finlândia	-0,47 (0,89)	-0,48 (0,86)	-1,33 (0,72)			
França	-0,02 (0,97)	-0,50 (0,67)	-0,40 (0,72)			
Holanda	0,82 (0,36)	1,12 (0,28)				
Itália	-0,74 (0,22)	-1,01 (0,19)				
Portugal	3,01 (0,14)	2,30 (0,50)	2,19 (0,50)	2,93 (0,31)		
Reino Unido	-4,00 (0,39)	-2,15 (0,64)	-2,56 (0,50)	-2,29 (0,83)	-2,54 (0,75)	-1,21 (0,94)

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.

Tabela A-14. Conselhos fiscais fracos – Efeito estimado por ano de pós-tratamento

Var. Dep.: SB	Efeito estimado por ano de pós-tratamento				
	(1)	(2)	(3)	(4)	(5)
África do Sul	0,21 (0,67)	-0,32 (0,47)			
Chile	-1,38 (0,18)	-3,04*** (0,00)			
Colômbia	-0,87 (0,14)	-0,21 (0,81)	1,53 (0,14)	2,20 (0,11)	
Espanha	-1,78 (0,53)	-1,01 (0,69)			
Eslováquia	2,03 (0,28)	0,93 (0,42)	-1,24 (0,64)	-0,47 (0,89)	0,89 (0,81)
Irlanda	-0,86 (0,75)	0,82 (0,89)	-1,00 (0,86)	0,01 (1,00)	
Luxemburgo	0,34 (0,86)	0,03 (1,00)			
Malta	4,78*** 0,00				
Peru	-0,47 (0,47)				

Notas: i) 36 países compõem o grupo de controle; ii) inferência bicaudal realizada com base na distribuição dos efeitos placebo; iii) p-valor padronizado pelo RMSPE. Erros-padrão em parênteses. ***, ** e * indicam nível de significância para 1%, 5% e 10%, respectivamente.