Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Policy Responses to Fiscal Restraints: A Difference-in-Discontinuities Design

Veronica Grembi, Tommaso Nannicini, Ugo Troiano

Working Paper n. 397

This Version: October, 2013

IGIER – Università Bocconi, Via Guglielmo Röntgen 1, 20136 Milano –Italy
http://www.igier.unibocconi.it

The opinions expressed in the working papers are those of the authors alone, and not those of the Institute, which takes non institutional policy position, nor those of CEPR, NBER or Università Bocconi.

Policy Responses to Fiscal Restraints: A Difference-in-Discontinuities Design*

Veronica Grembi
Copenhagen Business School

Tommaso Nannicini
Bocconi University, IGIER & IZA

Ugo Troiano
University of Michigan

This version, October 2013

Abstract

We evaluate the effect of relaxing fiscal rules on policy outcomes applying a quasi-experimental research design. In 1999 the Italian central government introduced fiscal rules aimed at imposing fiscal discipline on municipal governments, and in 2001 relaxed the rules for municipalities below 5,000 inhabitants. This shift allows us to implement a "difference-in-discontinuities" design by combining the before/after with the discontinuous policy variation. Our estimates show that relaxing fiscal rules triggers a substantial deficit bias, captured by a shift from a balanced budget to a deficit that amounts to 2 percent of the total budget. The deficit comes primarily from reduced revenues as unconstrained municipalities show lower real estate and income tax rates. Finally, we investigate the heterogeneity in policy responses across municipalities to provide new evidence on the costs and benefits of restricting fiscal policy. The impact is larger if the mayor can run for reelection, the number of political parties in the city council is higher, voters are older, and the performance of the mayor in providing public goods is lower, consistent with models of the political economy of fiscal adjustment.

JEL codes: C21, C23, H62, H72, H77.
Keywords: fiscal rules, local government finance, difference-in-discontinuities.

---

*A previous version of this paper circulated under the title "Do Fiscal Rules Matter? A Difference-in-Discontinuities Design." We thank Alberto Alesina, Robert Barro, Raj Chetty, and Larry Katz for detailed suggestions. We thank Daron Acemoglu, Philippe Aghion, Roland Benabou, Marianne Bertrand, David Cutler, John Friedman, Ed Glaeser, Josh Hausman, Nathan Hendren, Brian Knight, David Laibson, Amanda Pallais, Roberto Perotti, Torsten Persson, Per Pettersson-Lidbom, László Sándor, Frank Schilbach, Guido Tabellini, and seminar participants at Bank of Italy, Bologna University, Boston University, CELS 2012 Northwestern University, CEPR Public Policy Symposium 2012, Catholic University of Milan, CIFAR, CeSiFo, CREI, CSEF Naples, Erasmus University Rotterdam, Harvard Development, Labor and Macro Lunch, IGIER-Bocconi, IIES, NBER SI 2012, UCSD, University of Michigan, and USI Lugano for insightful comments. Errors are ours and follow a random walk. Financial support is gratefully acknowledged from Catholic University of Milan for Grembi; the ERC (Grant No. 230088) and Bocconi for Nannicini; Harvard Department of Economics and Multidisciplinary Program in Inequality & Social Policy, Bank of Italy and Bradley for Troiano. We also thank Giancarlo Verde, Carmine La Vita, Daniela De Marte (Italian Ministry of the Interior), and Salvatore Parlato (Ifel-ANCI) for help with data collection. *Corresponding author*: Ugo Troiano, University of Michigan, Department of Economics, Lorch Hall, 611 Tappan Avenue, Ann Arbor, MI 48109; e-mail: troiano@umich.edu.

1 Introduction

Can fiscal restraints imposed on local governments create incentives for reducing the accumulation of debt? The need for fiscal adjustment in the aftermath of the Great Recession has revived interest in fiscal rules aimed at disciplining the discretionary power of policy makers. Despite extensive research, the impact of fiscal restraints on debt accumulation and their effectiveness in reducing politically motivated deficits remain highly debated.[1] As the authors in the literature have acknowledged—see, for example, the discussion in Poterba (1996) or Alesina and Perotti (1996)—the search for a definitive conclusion is hampered by the potentially endogenous decision of whether to adopt fiscal rules or not.[2]

In this paper, we study the effect of relaxing fiscal restraints at the local government level.[3] We first show quasi-experimentally that fiscal rules do matter for restraining the accumulation of debt and that fiscal adjustment is concentrated in revenues. We then give evidence suggesting that the adjustment is driven by cities with more political distortions. We overcome previous data and identification limitations by using a novel identification design.

Our setting is Italy, where the central government set a target on deficit reduction for all municipal governments in 1999—the so called "Domestic Stability Pact," DSP henceforth—and relaxed it for municipalities below 5,000 inhabitants in 2001. This policy change allows us to combine two sources of variation, before/after 2001 and just below/above 5,000 inhabitants, and implement what we call a "difference-in-discontinuities" (or "diff-in-disc") design. It is important to note that a (standard) cross-sectional Regression Discontinuity (RD) design would not allow us to identify the effect of relaxing fiscal rules in this setting. The fact that there is another policy, started in the 1960s and still in place, according to which mayors of cities above 5,000 residents receive a higher salary, implies that analyzing the discontinuity at 5,000 residents in any given year would not identify the effect of the DSP. At the same time, a (standard) difference-in-differences design would require the strong assumption of parallel trend between small and large cities, which is not satisfied in our setting.

[1] As indicated by Drazen (2002) in a review article: "A key question (perhaps the key question) about fiscal rules is whether they have the effect of slowing the growth of deficits."

[2] Furthermore, there has been limited investigation on fiscal rules at the local government level, where forms of "hidden" public debt can grow and raise fears about the overall financial sustainability of a country. For a review of the current state of the literature, see Glaeser (2013).

[3] Many countries have adopted fiscal rules to discipline local governments, including Argentina, Austria, Brazil, Canada, China, Colombia, Czech Republic, Denmark, Italy, Mexico, Poland, Spain, Sweden, Turkey.

The intuition for our identification strategy is simple. The diff-in-disc estimator takes the difference between the cross-sectional discontinuity at 5,000 after 2001 (when both fiscal rules and the mayor's salary show a jump) and the cross-sectional discontinuity at 5,000 before 2001 (when only the mayor's salary shows a jump). We derive precise identifying assumptions under which this estimator can identify average treatment effects in a neighborhood of the 5,000 threshold, and propose specific diagnostics tools to assess the validity of the design. Assumptions are more local than those required for a difference-in-differences strategy, because a parallel trend between treated and untreated observations must be observed just at the threshold of interest and not in the whole sample. Moreover, the diff-in-disc assumptions may hold in empirical contexts—like ours—where the RD assumptions are not satisfied.

The main rule established by the DSP imposed a gradual reduction of the "fiscal gap," defined as the municipal deficit net of transfers and debt service. The rationale for the exemption of municipalities with less than 5,000 inhabitants in 2001 was to avoid burdening very small towns with onerous requirements, as they may be disadvantaged by economies of scale in managing the municipal government. The penalties put in place for not complying with the DSP included a cut in the annual transfers from the central government, a ban on new hires, and a cut on reimbursement and non-absenteeism bonuses. This means that, in the regulatory environment we study, there is a central authority that can collect standardized public accounts and enforce punishment for non-compliers.

A number of studies have argued that there are several reasons why fiscal rules might be ineffective in restraining fiscal policy (see, for example, the reviews by Alesina and Perotti, 1996, and Wyplosz, 2012).[4] The approval of the DSP itself was accompanied by widespread skepticism about its effectiveness, because Italy usually ranks last among OECD countries in ratings of law enforcement and government effectiveness (e.g., see Kaufmann, Kraay, and Mastruzzi, 2010). Thus, our results suggest that the lessons we draw from Italian cities on

[4]The fact that subnational policy makers have limited discretion in changing fiscal policy is the central reason for which fiscal rules on local governments might not work. Furthermore, most fiscal rules, including those we study, are not embedded in the Constitution. This implies that fiscal responsibility laws can be frequently changed and revised, and they might suffer from the same time inconsistency problem that characterizes fiscal policy. Additionally, Alesina and Perotti (1996) argue that lax enforcement is one of the reasons why fiscal rules might not work. This concern, however, is more relevant at the national level, since at the local level the central government can be a credible enforcement authority. Finally, rules usually target only some parts of the budget and this offers opportunities for policy makers to sidestep the rules by complying with them without changing the overall fiscal discipline—see Milesi-Ferretti (2003).

the effectiveness of fiscal restraints may extend to other regulatory environments where the fiscal authority setting the rules faces critical ex-ante commitment problems.

We divide our empirical analysis into three parts. First, we analyze the effects of relaxing fiscal restraints on the deficit, which is the main policy variable of interest, and on the fiscal gap, which is the main target of the law. We find that relaxing fiscal rules translates into a larger fiscal gap of about 40 to 60 percent over the course of the following four years. This large effect on the main target of the DSP has real consequences for policy outcomes, as unconstrained municipalities increase their deficit by 20 Euros per capita (2 percent of the total budget). The fact that we find an effect not only for the target of the DSP, but also for the main policy variable of interest not targeted by the law (debt accumulation) alleviates concerns arising from the possibility of creative accounting.

In the second part of our empirical analysis, we study the composition of the fiscal adjustment, by analyzing municipal financial reports and administrative data on municipal tax rates, which are set by the local policy makers. We find that unconstrained municipalities have statistically similar expenditure levels with respect to constrained municipalities, but have lower tax revenues. This difference can be partially explained by the additional finding that municipalities for which fiscal rules are relaxed set lower tax rates. The main tax rates decided by Italian cities are a real estate tax rate on home property (*Imposta Comunale sugli Immobili*, ICI), which provides almost 50 percent of municipal tax revenues, and a surcharge on the personal income tax (*Imposta sul Reddito delle Persone Fisiche*, IRPEF), which amounts for about 10 percent of municipal tax revenues. Cities for which fiscal rules are relaxed have both a lower real estate tax rate (by about 14 percent) and a lower income tax surcharge (by about 30 percent) after the policy shift.

Finally, in the third part of the analysis, we exploit the fact that our setup—that is, an exogenously imposed fiscal destabilization—can provide new evidence on when fiscal restraints matter the most. On the one hand, the optimal tax smoothing theory would suggest that countercyclical deficits can increase welfare by equalizing the distortionary cost of taxation across booms and recessions.[5] On the other hand, a persistent deficit bias might be the suboptimal result of the interplay between rational politicians, voters, and interest groups. Our empirical findings suggest that political factors play a first-order role in fiscal adjustments.

[5]See Barro (1974), Barro (1979), and Lucas and Stokey (1983).

We first compare municipalities where only two political parties are represented in the local legislative assembly (about half of the sample) versus municipalities with more parties. Our results show that relaxing fiscal rules increases the deficit only if more than two parties are seated in the assembly, which must approve the budget proposed by the mayor. This finding is consistent with models that explain deficits as the result of political fragmentation and of dynamic common pool (see Persson and Tabellini, 2000), and also with the cross-country evidence that coalition governments are associated with higher deficits (see Roubini and Sachs, 1989; Kontopoulos and Perotti, 1999).

We then study if the relaxation of fiscal rules is affected by whether the mayor faces a binding term limit or not. We find that the increase in deficit bias arises only for mayors who can be reelected. This result is consistent with models linking deficits to reelection incentives (see Aghion and Bolton, 1990) or to politicians' pandering to voters (see Maskin and Tirole, 2004). We also show that cities that increase the municipal deficit after the relaxation of fiscal rules have an older population. These results are consistent with the model of Song, Storesletten, and Zilibotti (2012), according to which young citizens have a disciplining role for fiscal policy because they internalize the future costs of present fiscal instability.

Finally, we move a first step in relating our findings to models that formalize the welfare costs and benefits of fiscal restraints. We show that the increase in deficit bias arises only for mayors that systematically under-provide the public goods promised to voters in the provisional budget. This evidence can be interpreted as consistent with the model of Besley (2007), to the extent to which politicians who persistently overpromise public goods are those who care more about reelection rather than social welfare.

Our results survive a large number of robustness checks. Among those, we use the introduction of the DSP for all municipalities in 1999 as a falsification test to show that our results are not driven by cities just below and just above 5,000 responding differently to the same set of fiscal rules. This test suggests that better paid (and, hence, better selected) mayors do not react differently to the relaxation of the DSP compared to the other mayors. This is also reassuring for the external validity of our results. Additionally, we repeat this falsification exercise by interacting our treatment variable with the heterogeneity dimensions discussed above, in order to show that the law did not bind differently across those subsamples. We also show that there is no difference in manipulative sorting around 5,000 between the pre-

treatment and the post-treatment period, and that municipalities just below and just above the threshold were on parallel trends in the pre-treatment period.

The paper proceeds as follows. Section 2 summarizes the relevant literature. Section 3 describes the Italian institutions. Section 4 lays out our identification strategy. Section 5 describes the data. Section 6 discusses the empirical results. We conclude with Section 7.

2 Related literature

This paper relates to two strands of literature. First, we contribute to the literature that has analyzed the effectiveness of fiscal rules.[6] Indeed, a number of empirical studies have evaluated whether fiscal rules are associated with sounder budget outcomes, reaching mixed conclusions. The evidence primarily comes from cross-country comparisons in specific regions, such as the European Union (see Hallerberg and Von Hagen, 1999; Debrun et al., 2008) or Latin America (see Alesina, Hausmann, Hommes, and Stein, 1996), and from local governments in a federal state, such as the U.S. (see Poterba, 1994; 1996).[7] While some studies find that fiscal rules do indeed result in lower budget imbalances (see Knight, 2000, who uses the difficulty of amending U.S. state constitutions as an instrument for analyzing the effect of supermajority requirements for tax increases on tax rates), others stress the reasons why they might not be effective (see Alesina and Perotti, 1999). We provide a quasi-experimental design where the effectiveness of fiscal rules is evaluated controlling for omitted factors that may affect previous results, such as the fact that more disciplined constituencies introduce tighter rules, or that (current and past) legal institutions are endogenous to cultural values.[8]

[6]For surveys, see Poterba and Von Hagen (1999), Rodden, Eskeland, and Litvack (2003), and Wyplosz (2012). For an extensive review on the types of rules and the main empirical evaluations of their impact, see IMF (2009). Balassone, Franco, and Zotteri (2004) review the literature on subnational fiscal rules in the European Union. As we focus on local governments, our results are particularly relevant for the literature that has emphasized that the implementation of subnational fiscal rules faces serious commitment problems, in the form of future overhaul, soft budget constraints, and lack of enforcement: see, among others, Eichengreen and von Hagen (1996), Braun and Tommasi (2004), Sutherland et al. (2005), and Ter-Minassian (2007).

[7]Studies on the U.S. also include Von Hagen (1991), Alt and Lowry (1994), Bayoumi and Eichengreen (1995), Bohn and Inman (1996), Alesina and Bayoumi (1996), Auerbach (2006), Fatas and Mihov (2006), Clemens (2012), and Clemens and Miran (2012).

[8]On the endogenous determination of laws, see Aghion, Alesina, and Trebbi (2004) and Givati and Troiano (2012). From a theoretical perspective, other authors analyze the welfare effect of fiscal restraints: Besley and Smart (2007) study limits on the size of government in a two-period agency model; Bassetto and Sargent (2006) study the welfare case for allowing the government to issue debt only to finance certain expenditures.

Secondly, we contribute to the large literature on the political economy of deficit determination, as we identify a set of politicians' and voters' characteristics associated with larger deficit bias.[9] From a normative perspective, it is not obvious whether tight rules such as a balanced budget requirement are optimal or not. As discussed above, restricting fiscal policy is not the first best in standard macroeconomic models without political distortions. However, restricting fiscal policy might become optimal when deficits are the suboptimal result of the interplay between rational politicians, voters, and interest groups (such as the aforementioned Alesina and Tabellini, 1990; Persson and Svensson, 1989; Aghion and Bolton, 1990; Besley, 2007; Battaglini and Coate, 2008; Song, Soresletten and Zilibotti, 2012).[10] These models deliver different predictions on the types of polities where one should expect the emergence of larger deficits. Our empirical findings shed new light on these dimensions.

3 Institutional framework

The Italian municipal government (*Comune*) is composed of a mayor (*Sindaco*), an executive committee (*Giunta*) appointed by the mayor, and an elected city council (*Consiglio Comunale*) that must endorse the annual budget proposed by the mayor. The mayor and the executive committee—whose members can be dismissed by the mayor at will—propose changes in fiscal policy, such as adjustments in the tax rates. Subsequently, the city council votes on the proposed changes. Since 1993, mayors are directly elected (with single round plurality rule in cities below 15,000 inhabitants) and face a two-term limit. Municipalities manage about 10 percent of total public expenditure and are in charge of a wide range of services, including water supply, waste management, municipal police, infrastructures, welfare, and housing. Only about 20 percent of revenues are local revenues.

After the European Union adopted its Stability and Growth Pact in 1997, some European countries—including Italy—adopted subnational fiscal rules to keep local governments accountable. In December 1998, the Italian annual budget law (*Legge Finanziaria*) for 1999 introduced a set of rules that constrained all municipalities in terms of fiscal discipline, the aforementioned Domestic Stability Pact or DSP (*Patto di Stabilità Interno*).[11]

[9] This literature has been reviewed by Alesina and Perotti (1999).

[10] Other political economy models on deficit determination include Tabellini and Alesina (1990), Lizzeri (1999), Azzimonti, Battaglini, and Coate (2008), and Yared (2010).

[11] See Law 23 December 1998, no. 448, article 28.

Municipal governments were constrained to keep the growth of their fiscal gap—defined as deficit, net of transfers and debt service—under tight control. The rationale for the exclusion of debt service and transfers in the definition of the DSP target is twofold. First, mayors are not held accountable for expenses on interest (which depend on previously contracted loans) and for revenues from transfers (which are not raised by the municipality). Second, these two items are the tools that the central government uses to enforce fiscal rules, reducing interest payments for compliers and cutting transfers for non-compliers. The punishment established for not complying with the DSP included the following penalties: (i) 5 percent cut in the annual transfers from the central government; (ii) ban on municipal hires; (iii) 30 percent cut on reimbursement and non-absenteeism bonuses for the employees of the municipal administration. Cities complying with the DSP, instead, benefited from a reduction of the expenses on interests for loans from the central government.[12]

The exact DSP rule constraining the fiscal gap changed from one year to another, but over our sample period it consisted in imposing a cap on the growth rate of the gap. Table 1 summarizes the evolution of the DSP over our sample period. The cap varies between a minimum of zero (no growth allowed) and a maximum of 3 percent, the benchmark being the fiscal gap two years before the actual budget year (this means that, for instance, the growth rate in 2004 is calculated with respect to the fiscal gap in 2002).

In evaluating the impact of the DSP on fiscal discipline, we therefore focus on the pattern of both deficit and fiscal gap. Constrained and unconstrained municipalities can accumulate debt, but if they run into fiscal distress they need to go through a special procedure of budget consolidation (*Piano di Risanamento*). One possible concern can be that relaxing fiscal rules induces expectations in our treated cities that they will be bailed out in case of situations of fiscal distress.[13] While we acknowledge the possibility that changes in fiscal restraints can always be confounded with changes in expectations, both legal and anecdotal evidence are consistent with the view that the Italian government made a substantial effort to keep

[12]In line with the law, we compute the fiscal gap with the formula: Fiscal Gap = (Total Expenditures - Debt Service) - (Total Revenues - Transfers). See the Appendix Table A1 for more details on the definition of policy outcomes. Unfortunately, the Ministry of the Interior does not release the list of municipalities that did not comply with the rule according to its records. As discussed in Section 6.1, we find suggestive evidence that the DSP penalties were to some extent enforced, as there is a correlation between non-compliance (as estimated in our data) and lower transfers (which are the main DSP enforcement mechanism).

[13]Italian cities can finance their debt through the emission of bonds (*Buoni Obbligazionari Comunali*) or with loans from a central administrative agency (*Cassa Depositi e Prestiti*) and from private banks.

expectations of bail outs as low as possible in the period of interest. In 2001, the Italian Constitution was revised to introduce a higher degree of fiscal decentralization while making bailouts unconstitutional.[14]

After 2001, all municipalities below 5,000 inhabitants were exempted by the DSP.[15] The motivation for this exemption was not made explicit by the central government, but it is probably linked to the goal of providing some relief to small municipalities in the presence of economies of scale in managing the municipal government.

Fiscal rules, however, are not the only policy varying with population size at 5,000. In particular, at this cutoff, there is a sharp increase in the wage received by the mayor and by the other members of the executive committee, based on a remuneration policy that has been in place since the early 1960s. Gagliarducci and Nannicini (2013) show that the wage increase at 5,000 attracts more educated individuals into politics and improves their performance once elected. Table 2 summarizes all the Italian policies on municipal governments relying on population thresholds over our sample period. Population size determines the size of the city council; the size of the executive committee; the electoral rule; and whether a municipality can have additional elective bodies at the neighborhood level. But only the DSP (after 2001) and the salary of local politicians display a discontinuity at the 5,000 threshold.

In 2002, regions with special autonomy (*Regioni a Statuto Speciale*) were allowed to set their own fiscal rules for municipal governments, and this is why we do not consider these regions in our study. Furthermore, since 2005 fiscal rules have been frequently changing from one year to another, shifting the population cutoff from 5,000 to 3,000 and back, and replacing the fiscal gap requirement with expenditure caps in some years. This is the reason why we focus our empirical evaluation on the period from 1997 to 2004.

In the next section, we explain how we exploit these unique Italian institutions to identify the effect of fiscal restraints on fiscal discipline.

[14]The new article 119 of the Italian Constitution specifically forbids the increase of governmental transfers to local governments in fiscal distress. Anecdotal evidence confirms a hard-line stance by the central government toward indebted municipalities. For instance, Taranto, a medium sized Italian city, declared bankruptcy in 2006; local newspapers reporting on the fiscal situation of the city (e.g., see *Taranto Sera*) stressed how the city had to undertake a multi-year repayment plan, without any help from the central government, and, after six years, almost half of the debt was still outstanding; public services and wage of public employees were suspended for some months after the bankruptcy, and local tax rates were significantly raised.

[15]See Law 23 December 2000, no. 388, article 53.

4 Difference-in-discontinuities design

4.1 Setup

Define $y_{it}(1)$ and $y_{it}(0)$ as the potential policy outcomes of municipality i at time t in the case of treatment ($d_{it} = 1$) and no treatment ($d_{it} = 0$), respectively. The observed outcome is therefore equal to: $y_{it} = d_{it}y_{it}(1) + (1 - d_{it})y_{it}(0)$. Because of the institutions described above, the treatment d_{it} coincides with "relaxing fiscal rules." If $t \geq t_0$, only municipalities below the population cutoff p_c are treated; the running variable p_{it} is set at the Census level. Formally, treatment assignment is equal to:

$$d_{it} = \begin{cases} 1 & \text{if } p_{it} < p_c \text{ and } t \geq t_0 \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Borrowing the notation from Hahn, Todd, and Van der Klaauw (2001), we define $z^- \equiv \lim_{p \to p_c^-} E[z_{it}|p_i = p, t \geq t_0]$ and $z^+ \equiv \lim_{p \to p_c^+} E[z_{it}|p_i = p, t \geq t_0]$, with $z = y(1), y(0), y$. Hahn, Todd, and Van der Klaauw (2001) derive precise conditions under which the cross-sectional RD estimator after t_0, defined as $\hat{\tau}_{RD} \equiv y^- - y^+$, identifies the average treatment effect at the threshold, $E[y_{it}(1) - y_{it}(0)|p = p_c]$. The identifying assumptions require d_{it} to be independent of $[y_{it}(1) - y_{it}(0)]$ conditional on p_i near p_c, and potential outcomes to be continuous at p_c, that is, $y(1)^- = y(1)^+$ and $y(0)^- = y(0)^+$.

In our setting, as it is often the case when different policies share the same threshold, the above continuity assumptions are not satisfied, because politician's wages also sharply change at p_c. Define: $\gamma_1 \equiv y(1)^- - y(1)^+ \neq 0$; $\gamma_0 \equiv y(0)^- - y(0)^+ \neq 0$. Here, γ_1 and γ_0 capture the effects of the confounding policy discontinuity on potential outcomes. It is straightforward to show that the cross-sectional RD estimator provides a biased estimate of the average treatment effects in a neighborhood of the threshold.[16]

[16]Specifically, the following relationships hold:

$$\hat{\tau}_{RD} \equiv y^- - y^+ = y(1)^- - y(0)^+ = [y(1)^- - y(0)^-] + \gamma_0, \tag{2}$$

$$\hat{\tau}_{RD} \equiv y^- - y^+ = y(1)^- - y(0)^+ = [y(1)^+ - y(0)^+] + \gamma_1, \tag{3}$$

where $[y(1)^- - y(0)^-]$ captures the average treatment effect for cities just below the threshold (treated after t_0), and $[y(1)^+ - y(0)^+]$ for cities just above the threshold (untreated after t_0). Only when $\gamma_1 = \gamma_0$, are the two estimands equal and represent the standard RD estimand, $E[y_{it}(1) - y_{it}(0)|p = p_c]$. In any case, $\hat{\tau}_{RD}$ always provides a biased estimate of the average treatment effects in a neighborhood of p_c.

4.2 Identification and diagnostics

We now show how to overcome the identification problem discussed above. Information on the pre-treatment period ($t < t_0$) allows us to remove the bias under local assumptions. Analogously to the post-treatment period, for the pre-treatment period we define: $\tilde{z}^- \equiv \lim_{p \to p_c^-} E[z_{it}|p_i = p, t < t_0]$ and $\tilde{z}^+ \equiv \lim_{p \to p_c^+} E[z_{it}|p_i = p, t < t_0]$, with $z = y(1), y(0), y$. To identify the causal effect of relaxing fiscal rules, we exploit both the (sharp) discontinuous variation at p_c and the outcome time variation after t_0:

$$\hat{\tau}_{DD} \equiv (y^- - y^+) - (\tilde{y}^- - \tilde{y}^+). \tag{4}$$

We call $\hat{\tau}_{DD}$ "difference-in-discontinuities" estimator (shortly, "diff-in-disc"), because it rests on the intuition of combining a difference-in-differences strategy and an RD design.

Alternative approaches in the literature have exploited the longitudinal nature of the data in an RD framework, such as the fixed-effect RD estimator in Pettersson-Lidbom (2012), the first-difference RD estimator in Lemieux and Milligan (2008), or the dynamic RD design in Cellini, Ferreira, and Rothstein (2011). All of these estimators, however, are different from ours. In their setups, treatment assignment changes over time and identification rests on within-unit variation, while in our case identification rests on the difference between two cross-sectional estimators. We are aware that other empirical studies implement some type of diff-in-disc strategy, but we now provide precise identification assumptions for this approach and propose diagnostics tools that directly stem from these assumptions.[17]

Assumption 1. *The effect of the confounding policy discontinuity on the potential outcome in the case of no treatment is constant over time:* $y(0)^- - y(0)^+ = \tilde{y}(0)^- - \tilde{y}(0)^+$.

This assumption can be interpreted from two perspectives. First, it is similar to the RD assumption that potential outcomes are continuous, as it states that the *difference* in $y_{it}(0)$ before and after t_0 is continuous at p_c. If the standard continuity assumption holds, as $\gamma_0 = 0$ at every t, Assumption 1 is also met and the diff-in-disc estimator could simply be used as

[17]For examples of diff-in-disc strategies, see Lalive (2008), Casas-Arce and Saiz (2011), Campa (2011), and Leonardi and Pica (2013). Our econometric strategy also relates to evaluation designs that exploit the comparison between different discontinuities across space, such as in different U.S. states (see Dickert-Conlin and Elder, 2010) or for politicians facing different term limits (see Gagliarducci and Nannicini, 2013).

a robustness check of the standard RD design.[18] But if the standard continuity assumption does not hold, as in our empirical setting, Assumption 1 might still be met. A first test to validate the plausibility of this (untestable) assumption coincides with checking whether any manipulation of the running variable changes (or arises) over time.[19]

From a second perspective, Assumption 1 requires the effect of the confounding policy discontinuity not to vary with time. In other words, it requires observations just below and just above p_c to be on a parallel trend. This is similar to the standard identifying assumption for difference-in-differences but is more local, as it must be met only in a neighborhood of the policy threshold.[20] To test for (local) parallel trend, in Section 6, we estimate the pattern of the discontinuities in y_{it} before t_0 and show that observations just below and just above p_c were not on differential trends before the policy shift.

Assumption 2. *The effect of the confounding policy discontinuity is the same in the case of treatment and no treatment:* $y(1)^- - y(1)^+ = y(0)^- - y(0)^+$.

This second (homogeneity) assumption states that there must be no interaction between the treatment and the confounding policy discontinuity, and is similar to the additivity condition in difference-in-differences. In our institutional setting, this assumption would be violated if mayors just below and above p_c, who are paid differently, reacted to fiscal rules in a different way. In Section 6, under the maintained hypothesis that Assumption 1 holds, we directly test this second condition exploiting the introduction of fiscal rules for all municipalities in 1999. In fact, if Assumption 2 holds, a falsification test implementing the diff-in-disc estimator in 1999 should deliver a zero effect.

It is straightforward to show that under the above assumptions, the diff-in-disc estimator identifies a (local) causal effect.

[18] Actually, Lalive (2008) uses a version of the diff-in-disc estimator (see footnote 31 in his article) exactly to evaluate the sensitivity of the results derived under a standard RD strategy.

[19] Specifically, in Section 6, we extend the cross-sectional test of continuity of the density at p_c (see McCrary, 2008) to test for the continuity of the *difference* in the densities before and after t_0. We also implement diff-in-disc estimations with time-invariant characteristics as outcomes, so as to indirectly test for changes in the pattern of manipulative sorting. As a further check in this direction, we include time-invariant characteristics and year fixed effects as covariates in the baseline diff-in-disc estimations; in the absence of manipulative sorting, point estimates are expected to remain similar and accuracy to increase.

[20] Indeed, in our empirical setting, the difference-in-differences assumption of parallel trend is unlikely to be satisfied (see the evidence discussed in the next section).

Proposition 1. *Under Assumption 1 and Assumption 2, the diff-in-disc estimator $\hat{\tau}_{DD}$ identifies the average treatment effect in a neighborhood of p_c: $E[y_{it}(1) - y_{it}(0)|p = p_c]$.*

Proof. Because of Assumption 1: $\hat{\tau}_{DD} \equiv (y^- - y^+) - (\tilde{y}^- - \tilde{y}^+) = (y(1)^- - y(0)^-) + (y(0)^- - y(0)^+) - (\tilde{y}(0)^- - \tilde{y}(0)^+) = y(1)^- - y(0)^-$. Because of Assumption 2: $y(1)^- - y(0)^- = y(1)^+ - y(0)^+ \equiv E[y_{it}(1) - y_{it}(0)|p = p_c]$. Therefore: $\hat{\tau}_{DD} = E[y_{it}(1) - y_{it}(0)|p = p_c]$.[21] □

4.3 Estimation

The diff-in-disc estimator can be implemented by estimating the boundary points of four regression functions of y_{it} on p_{it}: two on both sides of p_c, both before and after t_0. We borrow two different estimation methods from the RD literature for this purpose: local linear regression and spline polynomial approximation.[22]

The first method fits linear regression functions to the observations distributed within a distance h on either side of p_c, both before and after t_0. Formally, we restrict the sample to cities in the interval $p_{it} \in [p_c - h, p_c + h]$ and estimate the model:

$$y_{it} = \delta_0 + \delta_1 p^*_{it} + S_i(\gamma_0 + \gamma_1 p^*_{it}) + T_t[\alpha_0 + \alpha_1 p^*_{it} + S_i(\beta_0 + \beta_1 p^*_{it})] + \xi_{it}, \tag{5}$$

where S_i is a dummy for cities below 5,000 capturing treatment status, T_t an indicator for the post-treatment period, and $p^*_{it} = p_{it} - p_c$ the normalized population size. Standard errors are clustered at the city level. The coefficient β_0 is the diff-in-disc estimator and identifies the treatment effect of relaxing fiscal rules, as the treatment is $d_{it} = S_i \cdot T_t$. We present the robustness of our results to multiple bandwidths h, optimally computed first following the algorithm developed by Calonico, Cattaneo, and Titiunik (2013a, 2013b), and then implementing the cross-validation method proposed by Ludwig and Miller (2007).

The second method uses all observations and chooses a flexible functional form to fit the relationship between y_{it} and p_{it} on either side of p_c, both before and after t_0:

$$y_{it} = \sum_{k=0}^{q}(\delta_k p^{*k}_{it}) + S_i \sum_{k=0}^{q}(\gamma_k p^{*k}_{it}) + T_t\big[\sum_{k=0}^{q}(\alpha_k p^{*k}_{it}) + S_i \sum_{k=0}^{q}(\beta_k p^{*k}_{it})\big] + \xi_{it}. \tag{6}$$

[21] Note that even if Assumption 1 held but Assumption 2 did not, $\hat{\tau}_{DD}$ would still identify a causal effect: $[y(1)^- - y(0)^-]$. However, unlike the standard RD estimand in Proposition 1, which encompasses the entire neighborhood of p_c, this new estimand would only refer to (treated) observations just below p_c. Assumption 2 can thus be seen as a homogeneity condition that makes the (local) estimand more general.

[22] See Imbens and Lemieux (2008), Van der Klaauw (2008), and Lee and Lemieux (2010).

Standard errors are again clustered at the city level, and the coefficient β_0 is the diff-in-disc estimator identifying the treatment effect of relaxing fiscal rules. We present the robustness of our results to multiple orders of the polynomial approximation (q).

5 Data

We use administrative data from the Italian Ministry of the Interior (*Ministero dell'Interno*) containing information at the municipality level on the universe of municipal financial reports, municipal tax rates, electoral outcomes, and individual characteristics of the mayor. Based on the local nature of our diff-in-disc design, we restrict the sample to Italian municipalities between 3,500 and 7,000 inhabitants.[23] For the reason discussed in Section 3, we drop municipalities in regions with special autonomy. This leaves us with a final sample of 1,050 municipalities for a total of 6,300 observations from 1999 to 2004. Among them, 555 municipalities are treated after 2001 and 495 are in the control group. Our sample contains about 13 percent of all Italian municipalities and about 8 percent of the national population.

The population size comes either from the 1991 or from the 2001 Census. Because the relaxation of the DSP was decided in December 2000, it is very unlikely that municipalities had the time to influence their population and sort below the 5,000 threshold, and—on top of this—it is also unlikely that elected officials wanted to do that at the price of cutting their wage. In any case, in Section 6.2, we formally test for manipulative sorting below 5,000 before/after 2001 by comparing population size in the 1991 and 2001 Census.

The main variables of interest are the municipal financial report's categories. To measure fiscal discipline, we evaluate the deficit (total expenditures minus total revenues) and the fiscal gap (total expenditures minus total revenues, net of transfers and debt service), which is the target of the DSP. We divide expenditures into current outlays (including personnel expenditure), capital outlays (mostly investments), and debt service; and we divide revenues into municipal taxes, fees and tariffs, transfers from the central government, and other revenues. The main tax instruments decided by municipal governments are the real estate tax rate on home property (ICI), providing about 50 percent of their tax revenues, and the mu-

[23] We restrict the sample to the interval 3,500–7,000 to stay relatively far from the 3,000 threshold, where other policies change (see Table 2), and to balance the sample size on either side of the 5,000 threshold. All the results are robust to this interval choice, i.e., they are virtually unchanged for alternative choices, such as 3,250–6,750; 3,000–7,000; 3,500–6,500; 4,000–6,000; and 3,500–7,500 (available upon request).

nicipal surcharge on the personal income tax (IRPEF), amounting to about 10 percent of tax revenues.[24] See the Appendix Table A1 for precise definitions and data sources of all variables from the municipal financial reports.

One possible concern in evaluating the reaction of policies and tax instruments to fiscal rules might be that mayors have very little autonomy in adjusting local revenues or expenditure, but this is not the case for Italian municipalities. On the revenues side, over our sample period, mayors could vary ICI within a bracket from 0.4 to 0.7 percent of the legal home value, and the IRPEF surcharge within a bracket from 0 to 0.5 percent of taxable income.[25] And they were also free to set other local taxes (such as those on building rights or the occupation of public areas) or fees and tariffs for the services they provided (such as waste management or child care). Additionally, Italian towns are characterized by a sizable level of tax evasion, which the mayor can decide to fight.[26]

On the expenditure side, municipalities also have room for adjustment because about one third of expenditures are classified as not rigid (that is, not attributable to payroll expenses and debt service). For instance, one way to reduce expenditures without affecting the level of services is outsourcing (e.g., child care provided by private firms with more labor flexibility and lower costs although the financing remains public). Furthermore, Bandiera, Prat, and Valletti (2009) show how similar Italian municipalities can pay very differently for similar goods, and they interpret this as evidence of passive waste. This implies that, even if all current expenditures were rigid (and this is certainly not the case), mayors would still have the ability to reduce passive waste in order to adjust the fiscal gap.

Our dataset also contains time-invariant information on each municipality (geographic location, area size in km^2, sea level in meters), as well as time-varying information on the elected mayor (age, years of schooling, tenure in office, term limit), on the socio-economic environment (taxable income of resident inhabitants, age structure of the population), and on the political environment (number of political parties in the city council).

[24] Bordignon, Nannicini, and Tabellini (2013) also use ICI as the main policy tool of Italian municipalities.

[25] One additional concern can be that mayors comply with the rule by simply manipulating legal home value. However, legal home value is not determined or updated by mayors, as indicated by the *DPR* 22 December 1986, no. 917. Only in 2005, not in our sample, the Law 23 December 2005, no. 266 gave to municipalities some weak power of requesting the update of the assessed value of the real estate tax base.

[26] Casaburi and Troiano (2012) find that in 2007 over 2 million Italian buildings were not registered in the cadastral maps and thus were not part of the tax base for real estate and income tax.

Table 3 provides descriptive statistics on the main outcome variables (policy outcomes and tax instruments) for cities below and above 5,000 inhabitants. All variables are per capita and expressed in real terms (with 2009 as base year); tax rates are in percentage points. Municipalities below (above) 5,000 manage an annual budget equal to almost 1,041 (943) Euros per capita in terms of expenditures, and the deficit amounts to about 15 (11) Euros. Taxes are only slightly lower than 20 percent of total revenues and higher in municipalities above 5,000. The main tax rates on ICI and the IRPEF surcharge, however, are fairly similar for municipalities in the two groups.

As a benchmark, note that applying a difference-in-differences strategy to our dataset delivers the expected result: relaxing fiscal rules increases the deficit by 6.276 Euros per capita and the fiscal gap by 48.278 Euros per capita in a specification without municipality fixed effects, and by 5.279 and 16.669 with fixed effects, where both estimates on deficit are statistically significant at a 5 percent level and those on the fiscal gap at a 1 percent level. Those coefficients should be interpreted with caution, because their identification relies on assuming that small cities are a good counterfactual for larger cities. This is a strong assumption, which is likely to be violated because population affects the trends, and not only the levels, of public policies.[27]

In the next section, we discuss the results of our diff-in-disc design, which provides more credible evidence about the impact of relaxing fiscal rules on fiscal discipline because of the reasons discussed in Section 4.

6 Empirical results

6.1 Effect of relaxing fiscal rules on policy outcomes

Table 4 contains the main (diff-in-disc) estimation results. The main outcomes of interest are the two measures of fiscal discipline: deficit and fiscal gap (see panel A of the table). While the latter is the main target of the DSP, we believe that the former should be the real variable of policy interest. For each outcome variable, we show the robustness of the results to four

[27] In our sample, even if local trends are parallel around the 5,000 threshold (as shown in the next section), there is evidence that cities of different population size were on different global trends before the policy shift. From 1997 to 2000, the average deficit decreased by 30 percent in cities with a population above 5,000 inhabitants, while it increased by 53 percent in cities with a population below 5,000.

estimation methods: local linear regression as in equation (5), with two different types of optimal bandwidth; spline polynomial approximation as in equation (6), with two different orders of the polynomial (i.e., 3^{rd} and 4^{th}).[28]

The impact of relaxing rules on the deficit is positive and significant both in statistical and in economic terms. The DSP relaxation increases the deficit by about 20 Euros per capita with respect to a baseline situation of balanced budget. The deficit bias created by the relaxation of the DSP is also substantial from an economic point of view, as it amounts to about 2 percent of total expenditures. This effect is driven by a higher fiscal gap of about 40 to 60 percent, depending on the empirical specification. Both these effects are statistically significant at standard levels in all specifications, and the point estimates are stable in size across specifications with the exclusion of local linear regression with the larger optimal bandwidth (potentially more sensitive to the linearity assumption).

These estimation results on fiscal discipline are consistent with the descriptive graphs shown in Figure 1 where in the top graphs we draw scatters and (3^{rd}-order) polynomial fits of the differences between each post-2001 outcome value and each pre-2001 value. These graphs allow us to see whether those differences exhibit a discontinuity at the 5,000 threshold. We see that both variables measuring fiscal discipline exhibit a sharp jump when moving from the left to the right of the threshold in the whole sample. Furthermore, in the bottom graphs, we shed some light on the timing of the effect to provide evidence that high and low paid mayors were on parallel trends around the neighborhood of the 5,000 threshold. The evidence is consistent for both deficit and fiscal gap, as there is a change in the slope of the coefficients only after 2001. Most importantly, the cross-sectional RD estimates in 1999 and 2000 show that cities just below and just above 5,000 were on parallel trends before the policy shift.[29]

In panels B and C of Table 4, we assess whether the fiscal (de)stabilization takes place on the side of revenues or expenditure. While we cannot find a statistically significant impact on current outlays, capital outlays, and debt service, we find that tax revenues are lower

[28]Optimal bandwidths are calculated as either in Calonico, Cattaneo, and Titiunik (2013a, 2013b) or in Ludwig and Miller (2007). Results are robust to the use of predetermined bandwidths (i.e., 250, 500, 750, and 1,000) or additional orders of the polynomial (i.e., 2^{nd} and 5^{th}) and are available upon request. We use the two estimation methods that are standard in the RD literature. However, to address concerns about the sensitivity of our results to functional form assumptions, we also repeated the analysis implementing a simple t-test of the difference-in-discontinuities in closed intervals around the threshold (with intervals getting smaller and smaller) and we obtained similar results (available upon request).

[29]This robustness check also indirectly suggests that our policy was not anticipated in previous years.

by 20 to 45 percent in unconstrained municipalities (with respect to the average value of the control group and depending on the specification). Lower tax revenues are the result of lower tax rates decided by the municipal government (see panel D of Table 4). Cities for which fiscal rules are relaxed have a 14 percent lower real estate tax rate and a 30 percent lower income tax surcharge. Other revenues do not seem to be affected by the relaxation of fiscal restraints.[30] In Figure 3, we confirm that the (local) parallel trend assumption is also satisfied for all our other financial report's items.

Also on the side of revenues, central transfers seem to be higher for unconstrained municipalities, although point estimates are not always statistically significant. This result cannot explain the above impact of relaxing fiscal rules on fiscal discipline, because it goes in the opposite direction (that is, local governments running higher deficits receive larger transfers), and it is consistent with the design of the law, which allows the central government to cut transfers as an enforcement mechanism. This conjecture is consistent with our data. Although the Italian government did not release the official list of complying and non-complying municipalities, we can estimate compliance status in every year by applying the official rule summarized in Table 1 to our data. We find that complying municipalities amount to 68 percent of the total, and non-complying municipalities are also present around the 5,000 threshold, where the estimated compliance status shows a sharp discontinuity of about 40 percent.[31] Future transfers appear to be strongly correlated with compliance status: in a specification that controls for municipality and year fixed effects, central transfers are larger by about 10.329 Euros per capita (standard error, 3.303) for complying municipalities. This evidence is consistent with the institutional details discussed in Section 3, according to which central transfers are used as the main enforcement device of the DSP.

Figure 2 provides graphical evidence on the diff-in-disc jumps at the 5,000 threshold in the policy outcomes and tax instruments. Consistent with the estimation results, tax revenues,

[30]Other revenues include transfers from the European Union, other transfers, mortgages from administrative agencies, revenues coming from private properties owned by the municipality. Even if the standard errors for other revenues are bigger than the rest of our variables, visual inspection of the corresponding graph in Figure 2 reveals that standard errors are driven up by an outlier in this category. Repeating our analysis without this outlier consistently reduces the standard errors without affecting the other outcomes (results available upon request).

[31]Specifically, if we repeat our RD estimations using compliance status as a dependent variable, we obtain the following results for the local linear regression and the spline polynomial approximation specifications, respectively: 0.450 (standard error, 0.070); 0.443 (0.054); 0.448 (0.076); 0.436 (0.096).

ICI, and IRPEF show significant and negative jumps moving from just above to just below the 5,000 residents threshold. The graphs on expenditures confirm that these variables are reasonably stable across the threshold, providing support for the evidence suggesting that the big standard errors for expenditures are mainly due to outliers.

The previous results suggest that the fiscal adjustment mainly takes place through revenues. There are (at least) two possible explanations that can rationalize this finding. On the one hand, politicians might have a hard time convincing interest groups to cut expenditures, while taxpayers are more prone to internal free-riding and do not self-organize (see Olson, 1965). On the other hand, tax increases might be less salient than expenditure cuts for individuals (see Chetty, Looney, and Kroft, 2009) and thus they might be easier to adopt (or revert) for politicians.

6.2 Validity tests

As discussed in Section 4, the above estimation results rest on Assumption 1 and Assumption 2 for the identification of different average treatment effects in the neighborhood of the population threshold. In this section, we indirectly evaluate Assumption 1 by means of testing procedures aimed at detecting changes in manipulative sorting before/after 2001, and we directly test Assumption 2 in a falsification test that uses pre-treatment data.

The Appendix Figure A1 tests the null hypothesis of continuity of the *difference* in the density at 5,000 between the 1991 and the 2001 Census (top graph), by drawing both scatters and (3^{rd}-order) polynomial fits. If mayors were able to manipulate population size and sort below the threshold to avoid fiscal rules, our estimates would still suffer from the selection bias that was common in the previous empirical literature. However, in principle, there is very little room for differential manipulation between the two Censuses, because (i) the DSP is only enacted in December 2000, (ii) the Census is run independently by the National Statistical Office, so that false reporting should be ruled out, and (iii) mayors willing to sort below 5,000 to enjoy a relaxation of fiscal rules would pay the price of cutting their wage. Nevertheless, it might still be the case that some municipalities under financial stress tried to sort below 5,000 moving from the 1991 to the 2001 Census, by forcing some residents to leave or (more plausibly) not counter-reacting to population drops. Yet, the top graph in Figure A1 is reassuring about the absence of manipulation, as there is no jump in the difference

between the two densities. The point estimate from the spline polynomial approximation is equal to -0.078 (standard error, 0.114) and therefore is not statistically different from zero. For the sake of completeness, we also report the cross-sectional density tests for 1991 (bottom left) and 2001 (bottom right). Also there, there is no evidence of manipulation.[32]

Furthermore, in the Appendix Table A2, we check for the balancing of time-invariant characteristics by including covariates, together with year fixed effects, in the baseline diff-in-disc estimations; as expected, point estimates remain almost unchanged and accuracy increases. The Appendix Table A3 further evaluates the absence of manipulation. We implement diff-in-disc estimations with time-invariant characteristics (geographic location, area size, and sea level) as outcome variables, but we use changing population numbers: the 1991 Census before the treatment year, and the 2001 Census afterwards. This is meant to assess whether the fraction of cities with certain fixed characteristics just below or above 5,000 varies from 1991 to 2001. No time-invariant characteristics display a statistically significant jump.[33] We think that geographical location is a particularly interesting dimension here, because Italian geography is correlated with economic development, crime rates, labor market shirking, or political accountability (e.g., see Ichino and Maggi, 2000; Nannicini, Stella, Tabellini, and Troiano, 2013), and it could thus be associated with opportunistic manipulation too.

Based on this large amount of supporting evidence on Assumption 1, in Table 5 we directly test Assumption 2 under the maintained hypothesis that Assumption 1 holds. In particular, we check whether cities just below or just above the 5,000 threshold respond differently to fiscal rules. We use the introduction of the DSP in 1999 for *all* municipalities as an experiment to test for the absence of any differential response around 5,000. Specifically, we implement diff-in-disc estimations in the interval 1997–2000, using 1999–2000 as the post-treatment period and 1997–1998 as the pre-treatment period. All outcome variables are

[32]The 1991 point estimate is 0.068 (0.082); the 2001 point estimate is -0.010 (0.076).

[33]As an additional check, in the Appendix Table A4, we report balance tests of potentially endogenous characteristics. We implement diff-in-disc estimations with some (time-varying) economic and political characteristics of the municipality as outcome variables, using the 2001 Census population as the running variable as in the baseline specifications for the main policy outcomes. The time-varying characteristics we control for are the taxable income at the municipality level; the mayor's gender, years of schooling, and previous tenure in office; as well as the dimension of heterogeneous treatment effects we study in the next section (namely, term limit, number of parties, young cohorts, and speed of public good provision). These outcomes are potentially endogenous to the DSP, but detecting significant effects would disclose unexpected channels of adjustment through income, political selection, or public good delivery. This does not seem to be the case, as also these potentially endogenous variables are balanced around 5,000 before/after 2001.

perfectly balanced around the threshold before/after 1999, confirming the assumption that the DSP did not interact with the confounding wage discontinuity and did not bind differently across different sides of the population threshold.[34]

Finally, we perform a set of placebo tests to evaluate the possibility that our results arise from random chance rather than a causal relationship. In the Appendix Figures A2 and A3, in the spirit of DellaVigna and La Ferrara (2012), we implement—respectively for deficit and for fiscal gap—a set of diff-in-disc estimations at false population thresholds below and above 5,000 (namely, any point from 4,900 to 4,400 and from 5,100 to 5,600 in order to stay sufficiently away from the true policy threshold). At these false thresholds, we expect to find no systematic evidence of treatment effects similar to our baseline results. The two figures report the cumulative density function of these 1,000 placebo point estimates (using a specification with 3^{rd}-order spline polynomial), normalized with respect to our baseline point estimates for deficit and fiscal gap. This means, for instance, that a normalized coefficient of 80 stands for a placebo point estimate equal to 80 percent of the true baseline estimate at 5,000. The intuition here is that we should not observe too many normalized coefficients outside the interval from -100 to +100. Indeed, all of the placebo coefficients are below our estimated coefficients for both deficit and fiscal gap, and the cumulative density function of the normalized coefficients is much steeper around zero. Only 0.8 percent of the normalized placebo coefficients for the deficit are bigger than the true coefficient in absolute value, while none of them is so for the fiscal gap. On the whole, these placebo tests provide strong support for the robustness of our main results on fiscal discipline.

6.3 Political economy of fiscal adjustment

In this section, we exploit our research design to shed light on the political economy of the deficit. Evaluating the differential response of different politicians and voters to an exogenous (albeit local) variation in fiscal rules can identify important determinants of politically motivated deficits, and it provides new evidence about the costs and benefits of fiscal restraints. We start by looking at three political factors. First, we consider whether there are more than two parties in the city council—which must approve the budget proposed by the mayor—to

[34]The city of Romentino was an outlier due to a lucrative sale of land in 1998 and it was removed from the sample. Our results do not change with the inclusion of this city, with the exception of bigger standard errors for other revenues (available upon request).

capture political fragmentation and potential common pool problems. Second, we consider whether mayors face a binding term limit or not, because mayors in their second term have no reelection incentives and no personal stake in the city's budget for the following years. Third, we consider the age profile of citizens in our municipalities. We finally relate our findings to models that formalize the trade-off between reduced flexibility of fiscal policy and increased discipline on politically motivated deficit. Specifically, we consider the effort of the mayor in providing the public good promised to voters in the provisional budget.

Results for the first two variables are reported in Table 6, where we implement the baseline diff-in-disc estimations in split samples: (i) two parties in the city council (the median value) vs. more than two parties; (ii) binding vs. non-binding term limit. For each heterogeneity exercise, we report the diff-in-disc estimates in the two (split) subsamples, the difference between the two point estimates, as well as the Wald test p-value indicating whether this difference is statistically different from zero. We are aware that, while the point estimates in each subsample consistently identify local average treatment effects, the causal interpretation of their difference rests on an additional conditional independence assumption. This is why we also report a second Wald test p-value (with covariates) indicating whether this difference is robust to a specification including a full set of interactions with covariates at the municipality level (namely, the average taxable income; mayor's years of schooling; and whether the municipality is in the north of the country).[35] If this test is also statistically significant, it means that the differential impact of relaxing fiscal rules across our heterogeneity dimensions is not driven by those observable confounding city characteristics.[36]

First, we focus on political fragmentation. Political fragmentation generally arises when several agents have an active role in the allocation of the budget, each with its own constituency to please, and each with some weight in the final decision. There are two key determinants that affect how much a policy maker internalizes the costs of the demanded

[35]Note that we performed an additional robustness check for the heterogeneity dimensions that we study in this section. Specifically, we checked whether these variables are balanced at the 5,000 threshold. This is indeed what we find, suggesting that our strategy of splitting the sample is keeping the sample balanced in the neighborhood of 5,000 inhabitants as well (results available upon request).

[36]As a final robustness check on our heterogeneity analysis, we repeat the falsification test in 1999 for all the above dimensions. Specifically, in the Appendix Tables A5, A6, and A7 we implement the above heterogeneity diff-in-disc estimations in the interval 1997–2000, using 1999–2000 as the post-treatment period and 1997–1998 as the pre-treatment period. The fact that no effect and no difference are ever statistically significant means that municipalities around the threshold in different heterogeneity subsamples are not on differential trends before 2001. In other words, the DSP did not bind differently across those subsamples.

share of the budget: the number of decision makers participating in the bargaining process and the institutional rules determining the aggregation of preferences. Most empirical studies focus on the first determinant because of a lack of reliable proxies for the rules that determine the budget allocation across countries. We also follow this previous literature by focusing on the first determinant (see Kontopoulos and Perotti, 1999). However, one advantage of our setting is that we can safely assume that the rules that determine the allocation of the budget are constant around our threshold. The estimation results reported in Table 6 show that only municipalities with high political fragmentation react to the relaxation of the DSP, and this result is robust to controlling for covariates. This result is consistent with models that explain the deficit in terms of political fragmentation or dynamic common pool (see Persson and Tabellini, 2000) and also with the cross-country evidence that coalition governments lead to higher deficits (see Roubini and Sachs, 1989; Kontopoulos and Perotti, 1999).

Second, we focus on term limits, exploiting the fact that Italian mayors face a two-term limit.[37] Theoretical models suggest that the expectation of a future election can affect policies because politicians who plan to run again for office must please the voters sufficiently often to merit reelection (Barro, 1973; Banks and Sundaram, 1998). We find that the fiscal (de)stabilization induced by the relaxation of the fiscal restraints is driven by mayors without a binding term limit, although this result becomes borderline insignificant in some specifications that control for covariates. As mayors without term limits face both stronger reelection concerns and a higher expected probability that they (or their party) will remain in power, the above result provides more support for models linking deficits to reelection incentives (see Aghion and Bolton, 1990) or to politicians' pandering to voters (see Maskin and Tirole, 2004), rather than models viewing deficits as a way to tie the hands of future governments with different political preferences (see Alesina and Tabellini, 1990; Persson and Svensson, 1989; Tabellini and Alesina, 1990). Unfortunately, we are not able to provide further empirical evidence on strategic voting models because of the lack of a clear expected reelection probability outcome in our data.[38] We are also not able to rule out alternative channels that can rationalize our result on term limits, such as the possibility that political experience per se has an effect on how mayors react to the relaxation of fiscal rules, though it is encouraging

[37]It should be noted that: (i) municipalities do not vote at the same time, and (ii) the DSP was independent of local politics because it followed agreements between the European Union and its member countries.

[38]See Petterson-Lidbom (2001) for an empirical evaluation of strategic voting models.

that some of the results survive to controlling for the mayor's characteristics.[39]

In Table 7, which has the same structure of Table 6, we report one additional heterogeneity result along city characteristics. In particular, we implement the baseline diff-in-disc estimations in two separate subsamples: cities with a higher (i.e., above-median) fraction of young cohorts vs. the rest of the cities. Consistent with the model of Song, Storesletten, and Zilibotti (2012), the deficit increases after the relaxation of the rule only in cities with a larger proportion of young citizens. Song, Storesletten, and Zilibotti (2012) propose a dynamic politico-economic theory of fiscal policy for small open economies, and we view this model as particularly relevant for our setting because small cities are a reasonably good approximation of small open economies. The main intuition of their model is that younger citizens impose a disciplining effect on fiscal policy, because they internalize the future costs of a loose fiscal policy today. Both of the above predictions are borne out by our empirical findings.[40]

Finally, we move to relating our results to existing models that formalize the costs and benefits of restraining fiscal policy. Besley (2007) considers fiscal rules that increase the cost of issuing public debt in an environment characterized by two types of politicians, the ones who care about smoothing busyness cycle fluctuations and those who care about reelection. Rules that are not overly restrictive will bind more on bad politicians and increase voters' welfare. In Table 8, we split the sample based on the median of the speed of public good provision (measured as the ratio of paid outlays over the total outlays committed in the provisional budget). Our indicator of the speed of public good provision is calculated as a mayor-specific measure, averaged across all the years each mayor is in office.[41] The results show that the increase in deficit bias arises only for mayors that systematically under-provide the public goods promised to voters in the provisional budget. This evidence can be interpreted as consistent with a potential negative effect for voters following the relaxation of the rule, only

[39] Given that we can control for the selection of mayors (years of schooling), our findings provide additional support to the literature that focuses on the effect of term limit on political accountability and in-office performance (see Besley and Case, 1995; List and Sturm, 2006).

[40] Another testable prediction of their model is that deficit bias should be higher for cities where the mayor is affiliated with right-wing parties. While most of the mayors in our sample are not affiliated with parties that can be clearly mapped to the ideological spectrum, we also find evidence that is consistent with this prediction in the (small) subset of mayors affiliated to a political party (results available upon request).

[41] Rogoff (1990) argues that electoral incentives might distort fiscal policy because of the distorted incentives to over-provide public goods when it is more salient for voters.

to the extent to which politicians who consistently over-promise public goods are those who care more about reelection rather than social welfare.[42]

On the whole, the results discussed in this section suggest that fiscal restraints are more likely to bind for cities characterized by political failures, and that political economy factors play a first-order role in the process of fiscal adjustment.

7 Conclusion

Limiting the increase of public debt is a key policy issue in most economies. Fiscal rules are usually considered one of the potential solutions to public debt growth. In this paper we rely on a novel quasi-experimental design to show that fiscal rules enforced by a national government can be effective in causing a reduction of the accumulation of debt by local governments. Additionally, we are able to investigate the composition of fiscal adjustment and we show that unconstrained cities have lower tax rates and lower revenues following the relaxation of fiscal rules. We then link our results to existing theories of fiscal adjustment to provide new evidence about the costs and benefits of restraining fiscal policy. We show that deficit bias arises only where many parties are represented in the city council, mayors can run for reelection, there is a smaller fraction of young citizens, and mayors systematically underprovide the promised public good. These results suggest that fiscal restraints can be more effective when political distortions are larger.

We are aware that the enhanced internal validity of our evaluation design comes at the price of lower external validity, as is always the case in (local) econometric strategies based on policy discontinuities. However, our falsification test shows that mayors who are selected through a different mechanism—that is, a different salary—react in the same way to the introduction of fiscal rules. This suggests that our results can potentially apply to different institutional settings.

[42]In the Appendix Table A8, we check whether there are heterogeneous treatment effects for the composition of fiscal adjustment. We use as outcomes the three main tools that mayors can control to adjust their budget: current outlays, capital outlays, and local tax revenues. While we do not see systematic patterns from these results, we notice that in the subsample of mayors who do not face a binding term limit and in the one of mayors who are slower in providing public goods, part of the adjustment seem to happen on the expenditure side: those politicians react to the relaxation of the restraint by increasing capital outlays (although the result in the sample of mayors without binding term limit is not robust to the inclusion of covariates). Cities for which revenues are lower after the relaxation of the restraint are more likely to be ruled by a mayor supported by more than two parties and by mayors who are slower in providing public goods.

Our results raise a number of questions for further research. First, we show that fiscal rules, when accompanied by a proper enforcement mechanism, can be effective also in regulatory environments characterized by serious commitment issues such as the Italian case. Hence, fiscal rules might be useful in far more cases than those suggested by the conventional wisdom, and the optimal design of fiscal rules should take into account political incentives in the enforcement of the rules. Second, our results on the composition of fiscal adjustment suggest that stabilizing fiscal policy through revenues and through expenditures might not be politically equivalent. In our setting, politicians seem to be more prone to raising taxes rather than cutting expenditures to comply with an exogenous incentive for fiscal stability. Alesina and Ardagna (2012) raise the point that expenditure reductions are usually more expansionary than tax increases, regardless of the existence of a fiscal rule. Future research could explore the hypothesis that fiscal rules targeting expenditures may improve welfare compared to rules targeting deficit. Third, little is known about the political incentives that drive the choice between cutting expenditures and raising taxes to achieve stabilization. A rapidly growing literature in public economics have shown how the salience of tax changes affects behavioral responses (see Chetty, Looney and Kroft, 2009). It is an exciting direction for future research to investigate whether public good provision is subject to similar issues, and whether policy makers can exploit voters' behavioral biases in their favor.[43]

Our last set of empirical results implies that political incentives drive local government responses to exogenously imposed fiscal restraints. Since restricting tax smoothing is the main welfare cost of fiscal restraints identified by the macroeconomic literature, and since mechanisms for smoothing business cycle fluctuations, such as unemployment insurance, are often administered at the national level (as discussed by Gavin and Perotti, 1997), our results suggest that fiscal rules imposed on subnational governments might have limited welfare costs and significant benefits.

[43] A first attempt along these lines is represented by Bisin, Lizzeri, and Yariv (2011).

References

Aghion, P., Alesina, A., and F. Trebbi (2004): "Endogenous Political Institutions," *Quarterly Journal of Economics*, 119, 565- 612.

Aghion, P., and P. Bolton (1990): "Government Domestic Debt and the Risk of Default: a Political–Economic model of the Strategic Role of Debt," in R. Dornbusch and M. Draghi (eds.), *Public Debt Management: Theory and History*, Cambridge, UK: Cambridge University Press.

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

Alesina, A., and T. Bayoumi (1996): "The Costs and Benefits of Fiscal Rules: Evidence from U.S. States," NBER Working Paper 5614.

Alesina, A., and R. Perotti (1996): "Fiscal Discipline and the Budget Process," *American Economic Review Papers and Proceedings*, 86, 401–407.

Alesina, A., and R. Perotti (1999): "Budget Deficits and Budget Institutions," in J. Poterba and J. Von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, Chicago, IL: University of Chicago Press.

Alesina, A., and G. Tabellini (1990): "A Positive Theory of Fiscal Deficits and Government Debt," *Review of Economic Studies*, 57, 403–414.

Alesina, A., Hausmann, R., Hommes, R., and E. Stein (1999): "Budget Institutions and Fiscal Performance in Latin America," *Journal of Development Economics*, 59, 253–273.

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

Auerbach, A. (2006): "Budget Windows, Sunsets, and Fiscal Control," *Journal of Public Economics*, 90, 87–100.

Azzimonti, M., Battaglini, M., and S. Coate (2008): "Analyzing the Case for a Balanced Budget Amendment to the U.S. Constitution," mimeo.

Balassone, F., Franco, D., and S. Zotteri (2004): "Fiscal Rules for Sub-national Governments: Lessons from EMU Countries," in G. Kopits (ed.), *Rules-Based Fiscal Policy in Emerging Markets: Background, Analysis and Prospects*, Palgrave Macmillan.

Bandiera, O., Prat, A., and T. Valletti (2009): "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment," *American Economic Review*, 99, 1278–1308.

Banks, J., and R.K. Sundaram, (1998): "Optimal Retention in Agency Problems," *Journal of Economic Theory*, 82, 293–323.

Barro, R.J., (1973): "The Control of Politicians: An Economic Model," *Public Choice*, 14, 19–42.

Barro, R.J. (1974): "Are Government Bonds Net Wealth?" *Journal of Political Economy*, 82, 1095–1117

Barro, R.J. (1979): "On the Determination of the Public Debt," *Journal of Political Economy*, 87, 940–71.

Bassetto, M., and T.J. Sargent (2006): "Politics and Efficiency of Separating Capital and Ordinary Government Budgets," *Quarterly Journal of Economics*, 121, 1167–1210.

Battaglini, M., and S. Coate (2008): "A Dynamic Theory of Public Spending, Taxation, and Debt," *American Economic Review*, 98, 201–236.

Bayoumi, T., and B. Eichengreen (1995): "Restraining Yourself: The Implications of Fiscal Rules for Economic Stabilization," *IMF Staff Papers*, 42, 32–48.

Besley, T., and A. Case (1995): "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits," *Quarterly Journal of Economics*, 110, 769–798.

Besley, T. (2007): *Principled Agents? The Political Economy of Good Government*, Oxford: Oxford University Press.

Besley, T., and M. Smart (2007): "Fiscal restraints and voter welfare," *Journal of Public Economics*, 91, 755–773.

Bisin, A., A. Lizzeri, and L. Yariv (2011): "Government Policy with Time Inconsistent Voters," mimeo, New York University.

Bohn, H., and R. Inman (1996): "Balanced-Budget Rules and Public Deficits: Evidence from the U.S. States," *Carnegie-Rochester Conference Series on Public Policy*, 45, 13–76.

Bordignon, M., Nannicini, T., and G. Tabellini (2013): "Moderating Political Extremism: Single Round vs. Runoff Elections under Plurality Rule," IZA Discussion Paper 7561.

Braun, M., and M. Tommasi (2004): "Subnational Fiscal Rules: A Game Theoretic Approach," in Kopits, G. (ed.), *Rules-Based Fiscal policy in Emerging Markets: Background, Analysis and Prospects*, Palgrave Macmillan.

Calonico, S., Cattaneo, M. D., and R. Titiunik (2013a): "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs," mimeo, University of Michigan, Department of Economics.

Calonico, S., Cattaneo, M. D., and R. Titiunik (2013b): "Robust Data-Driven Inference in the Regression-Discontinuity Design," mimeo, University of Michigan, Department of Economics.

Campa, P. (2011): "Gender Quotas, Female Politicians and Public Expenditures: Quasi-Experimental Evidence," mimeo, Stockholm University.

Casaburi, L., and U. Troiano (2012): "Ghost-House Busters: The Electoral Response to a Large Anti Tax Evasion Program," mimeo, Harvard University.

Casas-Arce, P., and A. Saiz (2011): "Women and Power: Unwilling, Ineffective, or Held Back?" IZA Discussion Papers 5645.

Cellini, S.R., F. Ferreira, and J. Rothstein (2011): "The Value of School Facility Investments: Evidence from a Dynamic Regression Discontinuity Design," *Quarterly Journal of Economics*, 125, 215–261.

Chetty, R., Looney, A., and K. Kroft (2009): "Salience and Taxation: Theory and Evidence," *American Economic Review*, 99, 1145–77.

Clemens, J. (2012): "State Fiscal Adjustment During Times of Stress: Possible Causes of the Severity and Composition of Budget Cuts," mimeo.

Clemens, J., and S. Miran (2012): "The Effects of State Budget Cuts on Employment and Income," *American Economic Journal: Economic Policy*, 4, 46–68.

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*, 54, 297–362.

DellaVigna, S., and E. La Ferrara (2010): "Detecting Illegal Arms Trade," *American Economic Journal: Economic Policy*, 2(4), 26–57.

Dickert-Conlin, S., and T. Elder (2010): "Suburban Legend: School Cutoff Dates and the Timing of Births," *Economics of Education Review*, 29, 826–841.

Drazen, A. (2002): "Fiscal Rules From a Political Economy Perspective," paper prepared for the IMF–World Bank Conference *Rules-Based Fiscal Policy in Emerging Market Economies*, Oaxaca, Mexico, February 14–16.

Eichengreen, B., and von Hagen, J. (1996): "Fiscal Policy and Monetary Union: is There a Trade-Off Between Federalism and Budgetary Restrictions?," NBER Working Paper 5517.

Fatás, A., and I. Mihov (2006): "The Macroeconomic Effects of Fiscal Rules in the US States," *Journal of Public Economics*, 90, 101–117.

Gagliarducci, S., and T. Nannicini (2013): "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *Journal of the European Economic Association*, 11, 369–398.

Gavin, M., and R. Perotti (1997), "Fiscal policy in Latin America," *NBER Macroeconomics Annual* 12, 11–72.

Givati, Y., and U. Troiano (2012): "Law, Economics and Culture: Theory and Evidence from Maternity Leave Laws," *Journal of Law and Economics*, 5(2), 222–250.

Glaeser, E. (2013): "Urban Public Finance," in: A.J. Auerbach, R. Chetty, M. Feldstein, E. Saez (eds), *Handbook of Public Economics*, North-Holland.

Hahn, J., P. Todd, and W. Van der Klaauw (2001): "Identification and Estimation of Treatment Effects with Regression Discontinuity Design," *Econometrica*, 69, 201–209.

Hallerberg, M., and J. Von Hagen (1999): "Electoral Institutions, Cabinet Negotiations, and Budget Deficits in the EU," in J. Poterba and J. Von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, Chicago, IL: University of Chicago Press.

Ichino, A., and G. Maggi (2000): "Work Environment and Individual Background: Explaining Regional Shirking Differentials in a Large Italian Firm," *Quarterly Journal of Economics*, 15, 1057–90.

Imbens, G., and T. Lemieux (2008): "Regression Discontinuities Designs: A Guide to Practice," *Journal of Econometrics*, 142, 615-635.

International Monetary Fund (2009): *Anchoring Expectations for Sustainable Public Finances*, Washington DC.

Kaufmann, D., Kraay, A., and M. Mastruzzi (2010): "The Worldwide Governance Indicators: Methodology and Analytical Issues," World Bank Policy Research Working Paper 5430.

Knight, B.G. (2000): "Supermajority Voting Requirements for Tax Increases: Evidence from the States," *Journal of Public Economics*, 76, 41–67.

Kontopoulos, Y., and R. Perotti (1999): "Government Fragmentation and Fiscal Policy Outcomes: Evidence from OECD Countries," in J. Poterba and J. Von Hagen (eds.), *Fiscal Institutions and Fiscal Performance*, Chicago, IL: University of Chicago Press.

Lalive, R. (2008): "How Do Extended Benefits Affect Unemployment Duration: A Regression Discontinuity Approach," *Journal of Econometrics*, 142(2), 785–806.

Lemieux, T., and K. Milligan (2008): "Incentive Effects of Social Assistance: A Regression Discontinuity Approach," *Journal of Econometrics*, 142, 807–828.

Lee, D.S., and T. Lemieux (2010): "Regression Discontinuities Designs in Economics," *Journal of Economic Literature*, 48, 281–355.

Leonardi, M., and G. Pica (2013): "Who Pays for It? The Heterogeneous Wage Effects of Employment Protection Legislation," *Economic Journal*, forthcoming.

List, J.A., and D.M. Sturm (2006): "How Elections Matter: Theory and Evidence form Environmental Policy," *Quarterly Journal of Economics*, 121, 1249–1281.

Lizzeri, A. (1999): "Budget Deficits and Redistributive Politics," *Review of Economic Studies*, 66, 909–928.

Lucas, R., and N. Stokey (1983): "Optimal Fiscal and Monetary Policy in an Economy without Capital," *Journal of Monetary Economics*, 12, 55–94.

Ludwig, J., and D. L. Miller (2007): "Does Head Start Improve Children's Life Chances? Evidence from a Regression Discontinuity Design," *Quarterly Journal of Economics* 122(1): 159-208.

Maskin, E., and J. Tirole (2004): "The Politician and the Judge: Accountability in Government," *American Economic Review*, 94, 1034–1054.

McCrary, J. (2008): "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142, 698–714.

Milesi-Ferretti, G.M. (2003): "Good, Bad or Ugly? On the Effects of Fiscal Rules on Creative Accounting," *Journal of Public Economics*, 88, 377–394.

Nannicini, T., A. Stella, G. Tabellini, and U. Troiano (2013): "Social Capital and Political Accountability," *American Economic Journal: Economic Policy*, 5, 222–250.

Olson, M. (1965): *The Logic of Collective Action*, Cambridge, MA: Harvard University Press.

Persson, T., and G. Tabellini (2000): *Political Economics*, Cambridge, MA: MIT Press.

Persson, T., and L. Svensson (1989): "Why a Stubborn Conservative Would Run a Deficit: Policy with Time-Inconsistent Preferences," *Quarterly Journal of Economics*, 104(2), 325–345.

Pettersson-Lidbom, P. (2001): "An Empirical Investigation of the Strategic Use of Debt," *Journal of Political Economy*, 109, 570–584.

Pettersson-Lidbom, P. (2012): "Does the Size of the Legislature Affect the Size of Government: Evidence from Two Natural Experiments," *Journal of Public Economics*, 96, 269–278.

Poterba, J. (1994): "State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics," *Journal of Political Economy*, 102, 799–821.

Poterba, J. (1996): "Budget Institutions and Fiscal Policy in the U.S. States," *American Economic Review Papers and Proceedings*, 86, 395–400.

Poterba, J., and J. Von Hagen (eds) (1999): *Fiscal Institutions and Fiscal Performance*, Chicago, IL: University of Chicago Press.

Rodden, J.A., G.S. Eskeland, and J. Litvack (eds) (2003): *Fiscal Decentralization and the Challenge of Hard Budget Constraints*, Cambridge, MA: MIT Press.

Rogoff, K. (1990): "Equilibrium Political Budget Cycles," *American Economic Review*, 80, 21–36.

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

Song, Z., K. Storesletten, and F. Zilibotti (2012): "Rotten Parents and Disciplined Children: A Dynamic Politico-Economic Theory of Public Expenditure and Debt," *Econometrica*, 80(6), 2785–2803.

Sutherland, D., R. Price, and I. Joumard (2005): "Fiscal Rules for Sub-Central Governments: Design and Impact," OECD Economics Department Working Paper 465.

Tabellini, G., and A. Alesina (1990): "Voting on the Budget Deficit," *American Economic Review*, 80, 37–49.

Ter-Minassian, T. (2007): "Fiscal Rules for Sub-national Governments: Can They Promote Fiscal Discipline?" *OECD Journal on Budgeting*, 6(3).

Van der Klaauw, W. (2008): "Regression-Discontinuity Analysis: A Survey of Recent Developments in Economics," *Labour*, 22, 219–245.

Von Hagen, J. (1991): "A Note on the Empirical Effectiveness of Formal Fiscal Restraints," *Journal of Public Economics*, 44, 199–210.

Wyplosz, C. (2012), "Fiscal Rules: Theoretical Issues and Historical Experiences," NBER Working Paper 17884.

Yared, P. (2010): "Politicians, Taxes and Debt," *Review of Economic Studies*, 77, 806–40.

Tables and figures

Table 1: The rules of the Domestic Stability Pact (DSP)

Year	Target of the DSP rules	Covered municipalities
1997	None	All
1998	None	All
1999	Fiscal gap: zero growth	All
2000	Fiscal gap: zero growth	All
2001	Fiscal gap: max 3% growth	Above 5,000
2002	Fiscal gap: max 2.5% growth	Above 5,000
2003	Fiscal gap: zero growth	Above 5,000
2004	Fiscal gap: zero growth	Above 5,000

Notes. The *Domestic Stability Pact* is a set of fiscal rules imposed by the central government to discipline the fiscal management of local governments. The main target is the *Fiscal gap* (see Appendix Table A4 for details). The growth of the fiscal gap with respect to its value two years before is constrained to be either zero or below 2.5%/3% depending on the year of the DSP. Legislative sources: annual national budget law (*Legge Finanziaria*) from 1999 to 2004.

Table 2: Legislative thresholds for Italian municipalities, 1997–2004

Population	Wage of mayor	Wage of executive committee	Size of executive committee	Size of city council	Electoral rule
Below 1,000	1,291	15%	4	12	single
1,000-3,000	1,446	20%	4	12	single
3,000-5,000	2,169	20%	4	16	single
5,000-10,000	2,789	50%	4	16	single
10,000-15,000	3,099	55%	6	20	single
15,000-30,000	3,099	55%	6	20	runoff
30,000-50,000	3,460	55%	6	30	runoff
50,000-100,000	4,132	75%	6	30	runoff
100,000-250,000	5,010	75%	10	40	runoff
250,000-500,000	5,784	75%	12	46	runoff
Above 500,000	7,798	75%	14-16	50-60	runoff

Notes. Policies varying at different legislative thresholds in the period 1999–2004. *Population* is the number of resident inhabitants as measured by the last available Census. *Wage of mayor* and *Wage of executive committee* refer to the monthly gross wage of the mayor and the members of the executive committee, respectively; the latter is expressed as a percentage of the former, which refers to 2000 and is measured in Euros. *Size of executive committee* is the maximum allowed number of executives appointed by the mayor. *Size of city council* is the number of seats in the city council. The wage thresholds at 1,000 and 10,000 were introduced in 2000; all of the other thresholds date back to 1960. Since 1993, the *Electoral rule* for the mayor is plurality with either single round or runoff.

Table 3: Outcome variables, descriptive statistics

	Municipalities above 5,000	Municipalities below 5,000
A. Fiscal discipline		
Deficit	11.080	15.457
Fiscal gap	170.724	208.624
B. Expenditures		
Current outlays	475.312	502.181
Capital outlays	438.838	508.794
Debt service	29.139	30.107
C. Revenues		
Taxes	194.887	175.825
Fees & tariffs	56.601	58.938
Central transfers	188.783	223.274
Other revenues	491.938	567.589
D. Tax instruments		
Real estate tax rate	0.587	0.576
Income tax surcharge	0.309	0.309
Obs.	2,970	3,330

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. The average values of per-capita policy outcomes are in 2009 Euros. The real estate tax rate and the income tax surcharge are in percentage points; the former can vary from 0.4 to 0.7 percent; the latter can vary from 0 to 0.5 percent.

Table 4: The effect of relaxing fiscal rules, diff-in-disc estimates

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
A. Fiscal discipline				
Deficit	17.099**	9.454**	21.449**	25.109**
	(7.945)	(4.343)	(9.485)	(12.756)
Obs.	2,266	5,858	6,300	6,300
h	555	1498		
Fiscal gap	81.410**	48.469**	102.202***	108.128**
	(33.890)	(23.315)	(38.463)	(48.372)
Obs.	1,926	3,438	6,300	6,300
h	471	833		
B. Expenditures				
Current outlays	-62.101	-10.665	-32.366	-60.520
	(63.586)	(32.756)	(55.631)	(77.460)
Obs.	1,656	4,112	6,300	6,300
h	407	979		
Capital outlays	72.127	-4.221	91.321	202.679
	(100.513)	(83.336)	(103.145)	(140.024)
Obs.	1,500	3,974	6,300	6,300
h	370	944		
Debt service	-1.876	-2.096	-2.338	-2.375
	(8.490)	(3.587)	(6.631)	(9.675)
Obs.	1,570	4,908	6,300	6,300
h	388	1202		

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on policy outcomes and tax instruments below 5,000 after 2001. Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. All policy outcomes are per capita and in 2009 Euros. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 4: The effect of relaxing fiscal rules, diff-in-disc estimates (cont'd)

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
C. Revenues				
Taxes	-66.85**	-34.748*	-57.028**	-85.077**
	(31.828)	(20.166)	(27.193)	(35.162)
Obs.	1,352	2,810	6,300	6,300
h	334	684		
Fees & tariffs	-5.494	1.413	1.173	-4.051
	(11.058)	(7.199)	(10.601)	(13.910)
Obs.	1,862	3,238	6,300	6,300
h	453	795		
Central transfers	34.158	32.938	78.414**	80.644*
	(28.908)	(21.721)	(35.334)	(43.517)
Obs.	2,216	3,438	6,300	6,300
h	537	833		
Other revenues	3.127	-81.308	12.608	123.159
	(128.275)	(62.926)	(118.470)	(165.666)
Obs.	1,420	5,858	6,300	6,300
h	351	1498		
D. Tax instruments				
Real estate tax rate	-0.055**	-0.027*	-0.056**	-0.060*
	(0.026)	(0.016)	(0.026)	(0.033)
Obs.	1,636	3,806	6,300	6,300
h	402	907		
Income tax surcharge	-0.045	-0.044*	-0.058	-0.111**
	(0.042)	(0.026)	(0.041)	(0.051)
Obs.	1,208	2,594	4,588	4,588
h	409	871		

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on policy outcomes and tax instruments below 5,000 after 2001. Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. All policy outcomes are per capita and in 2009 Euros. Tax instruments are in percentage points. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5: Falsification test in 1999

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
A. Fiscal discipline				
Deficit	4.112	2.640	0.942	-1.520
	(7.158)	(5.120)	(8.997)	(10.673)
Obs.	1,400	3,816	4,176	4,176
h	571	1,498		
Fiscal gap	8.492	0.433	12.440	-2.693
	(11.070)	(7.558)	(11.882)	(15.294)
Obs.	981	2,178	3,132	3,132
h	529	1,132		
B. Expenditures				
Current outlays	0.764	-2.325	-11.384	-6.307
	(10.374)	(9.959)	(10.605)	(11.584)
Obs.	660	774	3,132	3,132
h	338	401		
Capital outlays	-2.041	-34.087	-13.273	-93.598
	(57.701)	(36.921)	(59.126)	(98.401)
Obs.	771	2,109	3,132	3,132
h	400	1,092		
Debt service	-0.053	-0.683	-0.786	0.368
	(1.325)	(0.882)	(1.278)	(1.591)
Obs.	762	1,968	3,132	3,132
h	393	1,022		

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1997 and 2000. Diff-in-disc estimates of the (false) impact of introducing fiscal rules on policy outcomes below 5,000 after 1999 (when no discontinuity was introduced by the DSP; see Table 1). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. All policy outcomes are per capita and in 2009 Euros. *Fiscal gap*, *Current outlays*, *Capital outlays*, and *Debt service* are not available in 1997; for these variables the observations in the four estimations, respectively, are: 945; 1,389; 3,135; 3,135. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5: Falsification test in 1999 (cont'd)

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
C. Revenues				
Taxes	-4.094	-6.709	-3.411	-2.401
	(4.557)	(5.116)	(4.270)	(4.923)
Obs.	856	716	4,176	4,176
h	330	281		
Fees & tariffs	-0.064	1.536	-0.654	0.241
	(3.735)	(3.174)	(3.428)	(3.490)
Obs.	764	1,012	4,176	4,176
h	298	392		
Central transfers	-0.736	4.026	-3.298	-11.149
	(6.260)	(5.842)	(7.187)	(9.013)
Obs.	1,316	2,900	4,176	4,176
h	531	1,129		
Other revenues	14.456	-10.373	27.041	-62.725
	(43.546)	(30.651)	(43.639)	(86.491)
Obs.	840	2,388	4,176	4,176
h	323	945		
D. Tax instruments				
Real estate tax rate	0.000	0.004	0.002	0.002
	(0.009)	(0.006)	(0.009)	(0.011)
Obs.	1,008	2,900	4,176	4,176
h	389	1,129		

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1997 and 2000. Diff-in-disc estimates of the (false) impact of introducing fiscal rules on policy outcomes below 5,000 after 1999 (when no discontinuity was introduced by the DSP; see Table 1). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. All policy outcomes are per capita and in 2009 Euros. The real estate tax rate is in percentage points (the income tax surcharge is not available for this test because it was introduced in 1999). Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 6: The political economy of deficit bias, part I

	LLR *CCT*	LLR *CV*	Spline poly 3^{rd}	Spline poly 4^{th}
		With two parties or less:		
Deficit	2.081	0.958	3.278	2.993
	(9.104)	(5.252)	(10.369)	(12.379)
Obs.	1,273	3,347	3,584	3,584
		With more than two parties:		
Deficit	46.160***	21.464***	50.869***	69.627**
	(15.645)	(6.934)	(19.446)	(27.413)
Obs.	876	2,505	2,716	2,716
Difference between the two subsamples	44.079	20.506	47.591	66.634
Wald test p-value without covariates	0.032	0.118	0.022	0.044
Wald test p-value with covariates	0.045	0.719	0.033	0.049
		With binding term limit:		
Deficit	7.213	1.937	4.017	5.598
	(7.588)	(4.715)	(9.599)	(12.083)
Obs.	1,203	2,588	2,780	2,780
		Without binding term limit:		
Deficit	30.123**	15.229**	33.047**	36.640*
	(12.800)	(6.809)	(14.471)	(20.332)
Obs.	1,211	3,261	3,520	3,520
Difference between the two subsamples	22.910	13.292	29.030	31.042
Wald test p-value without covariates	0.039	0.110	0.090	0.189
Wald test p-value with covariates	0.078	0.117	0.076	0.227

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on fiscal discipline below 5,000 after 2001 in different subsamples (that is, above vs. below median number of parties; binding vs. non-binding term limit). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—*CCT*—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—*CV*—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* evaluates whether the estimates are statistically different in the two subsamples also controlling for a full set of interactions between the above specifications and appropriate covariates, such as: average taxable income; mayor's years of schooling; and whether the municipality is in the North. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 7: The political economy of deficit bias, part II

	LLR *CCT*	LLR *CV*	Spline poly 3^{rd}	Spline poly 4^{th}
		Young cohorts above median:		
Deficit	2.033	1.353	7.094	4.151
	(10.390)	(6.193)	(11.067)	(14.635)
Obs.	1,121	2,960	3,224	3,224
		Young cohorts below median:		
Deficit	32.824***	18.744***	36.913***	49.481***
	(11.713)	(5.766)	(14.255)	(18.795)
Obs.	1,086	2,884	3,076	3,076
Difference between the two subsamples	30.791	17.391	29.819	45.330
Wald test p-value without covariates	0.031	0.035	0.082	0.041
Wald test p-value with covariates	0.037	0.068	0.152	0.048

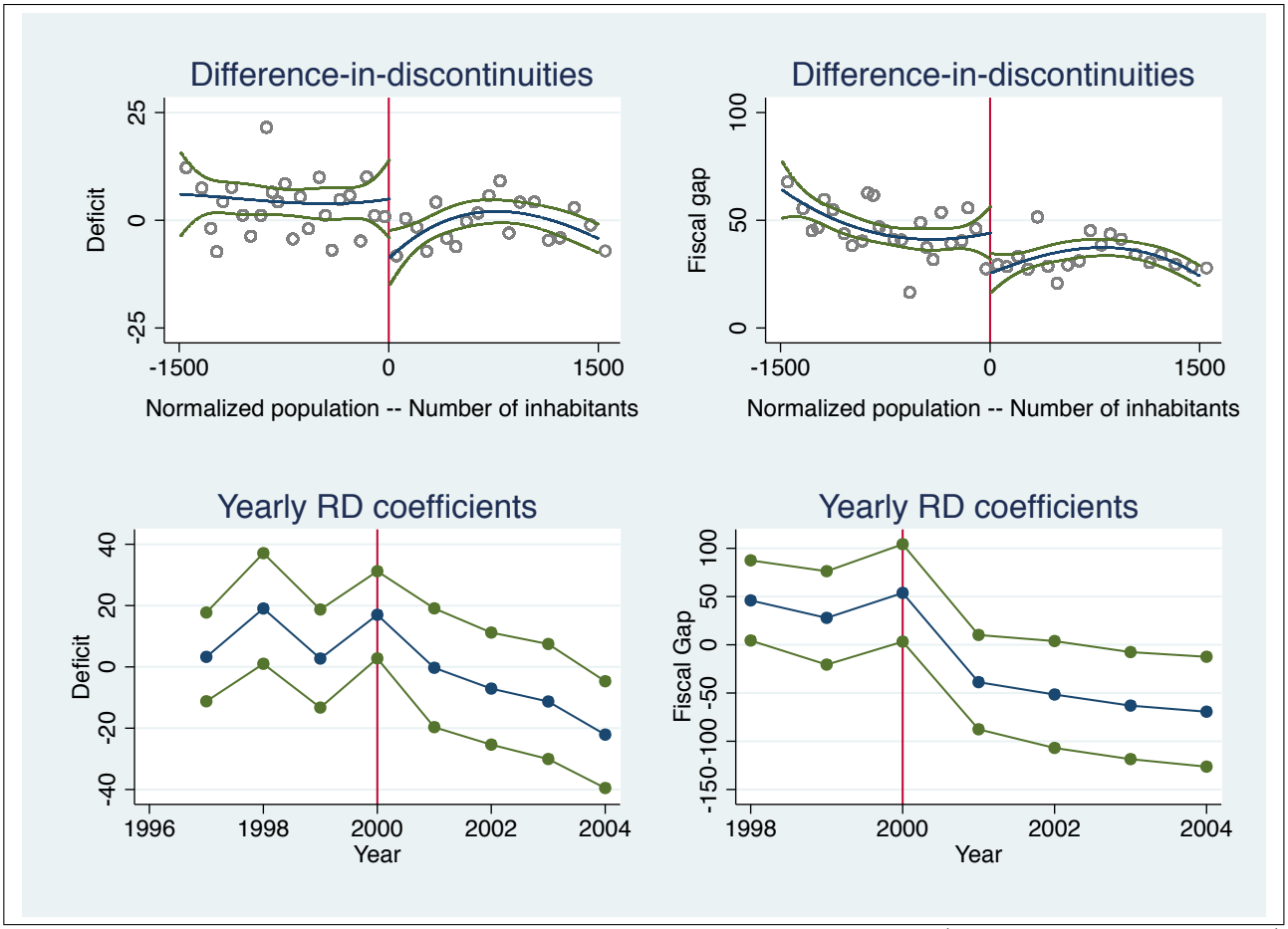
Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on fiscal discipline below 5,000 after 2001 in different subsamples (that is, above vs. below median percentage of young cohorts). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—*CCT*—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—*CV*—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* evaluates whether the estimates are statistically different in the two subsamples also controlling for a full set of interactions between the above specifications and appropriate covariates, such as: average taxable income; mayor's years of schooling; and whether the municipality is in the North. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 8: Fiscal restraints and budget management

	LLR *CCT*	LLR *CV*	Spline poly 3^{rd}	Spline poly 4^{th}
	Speed of public good provision above median:			
Deficit	-0.580	-0.604	4.433	3.717
	(9.805)	(6.190)	(11.602)	(16.051)
Obs.	1,267	2,951	3,153	3,153
	Speed of public good provision below median:			
Deficit	33.433***	18.975***	38.184**	47.452**
	(11.559)	(6.035)	(14.811)	(20.275)
Obs.	1,163	2,897	3,147	3,147
Difference between the two subsamples	34.013	19.579	33.751	43.735
Wald test p-value without covariates	0.030	0.019	0.065	0.086
Wald test p-value with covariates	0.073	0.016	0.052	0.057

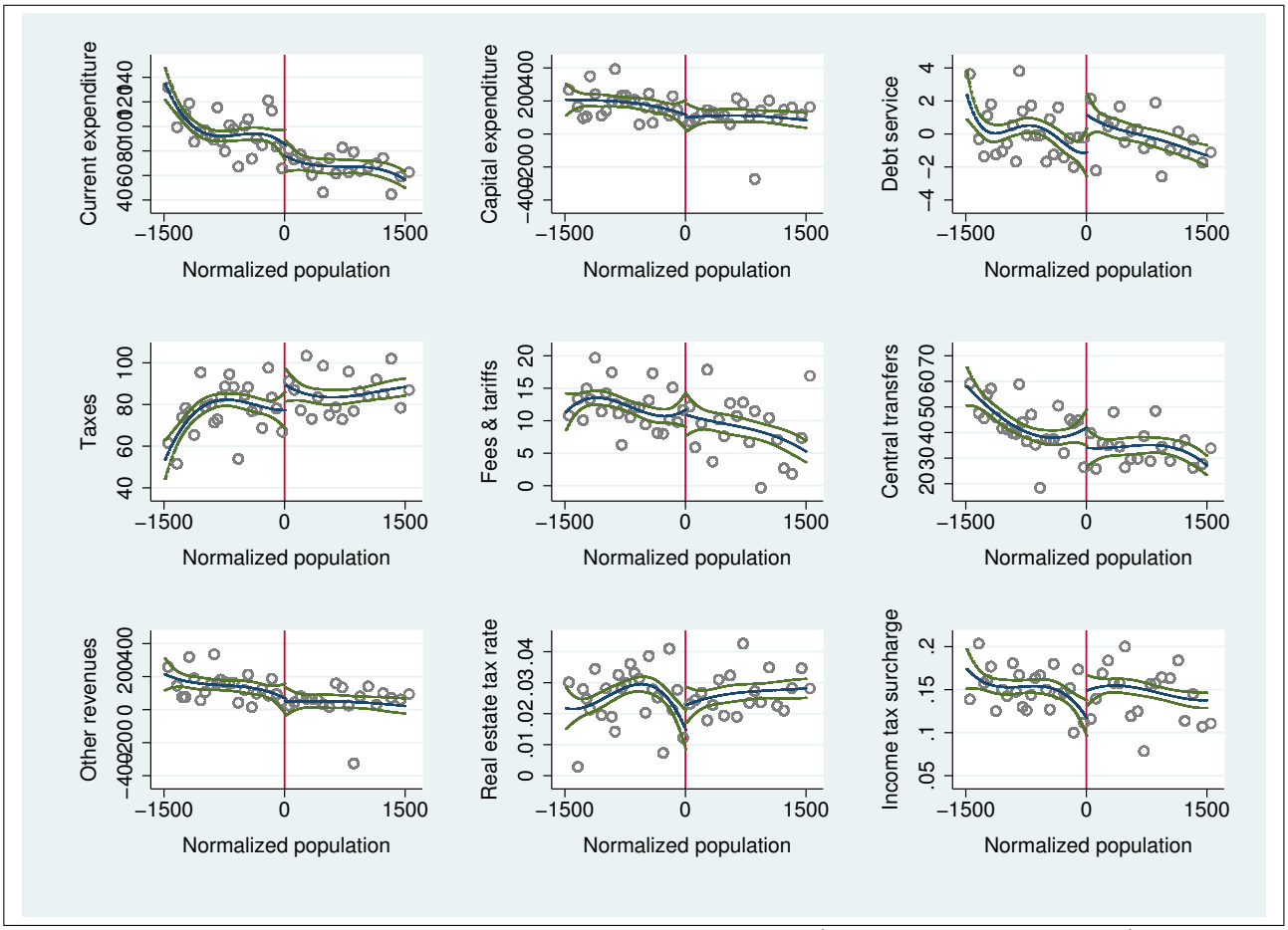
Notes. Municipalities between 3,500 and 7,000 inhabitants; between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on fiscal discipline below 5,000 after 2001 in different subsamples (that is, above vs. below median speed of public good provision). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—*CCT*—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—*CV*—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* evaluates whether the estimates are statistically different in the two subsamples also controlling for a full set of interactions between the above specifications and appropriate covariates, such as: average taxable income; mayor's years of schooling; and whether the municipality is in the North. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 1: Difference-in-discontinuities for deficit and fiscal gap



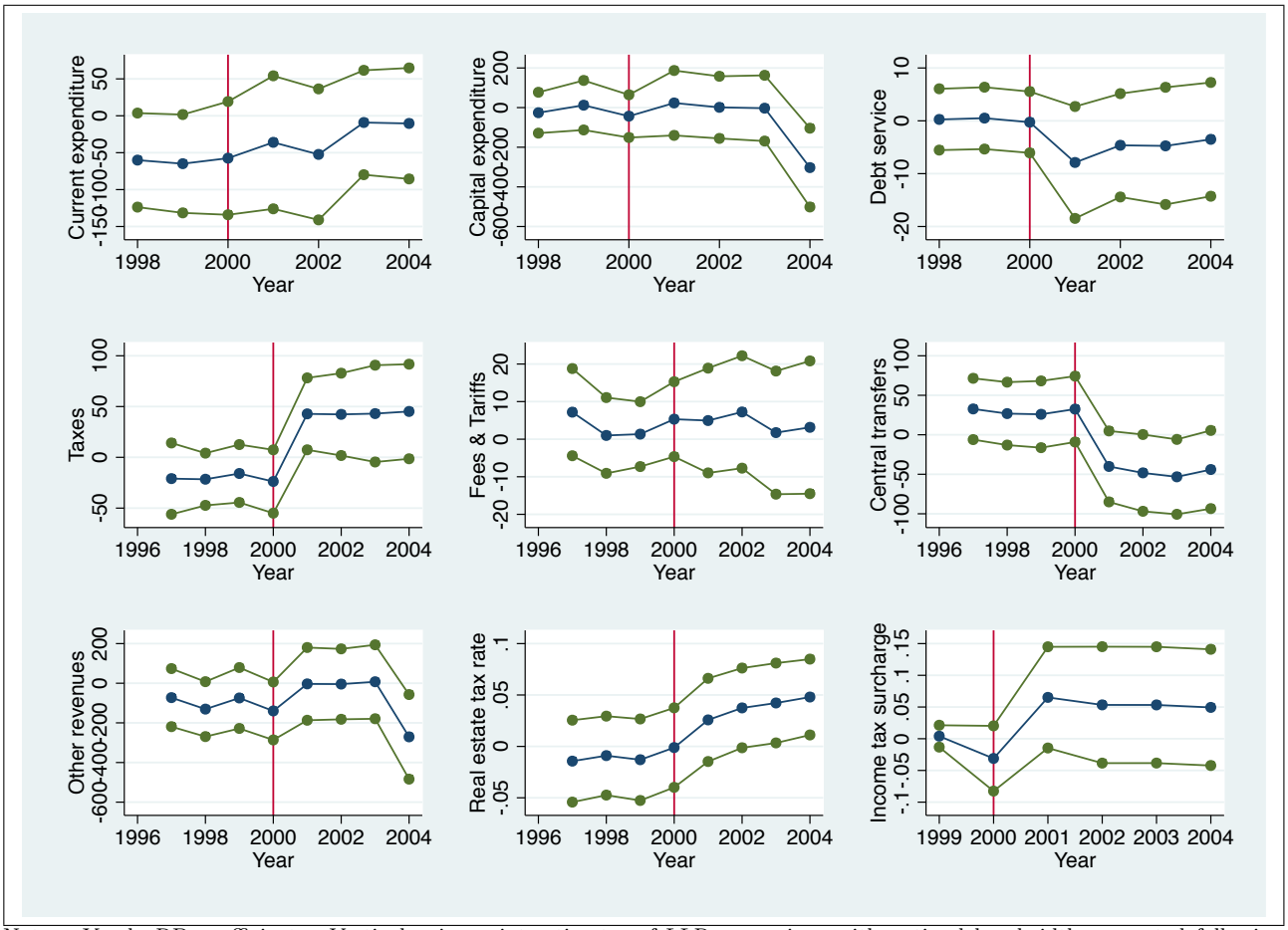
Notes. Top graphs: difference-in-discontinuities. Vertical axis: difference of each post-rule (i.e., 2001, 2002, 2003, and 2004) outcome value and each pre-rule (i.e., 1999 and 2000) outcome value. Horizontal axis: actual population size minus 5,000. The central line is a spline 3^{rd}-order polynomial fit; the lateral lines represent the 95% confidence interval. Scatter points are averaged over intervals of 50 inhabitants. Bottom graphs: yearly RD coefficients. Vertical axis: point estimates of LLR regressions with optimal bandwidth computed following Calonico, Cattaneo, and Titiunik (2013a, 2013b). Horizontal axis: year. The central line is the point estimate; the lateral lines represent the 95% confidence interval.

Figure 2: Difference-in-discontinuities for policy outcomes and tax instruments



Top graphs: difference-in-discontinuities. Vertical axis: difference of each post-rule (i.e., 2001, 2002, 2003, and 2004) outcome value and each pre-rule (i.e., 1999 and 2000) outcome value. Horizontal axis: actual population size minus 5,000. The central line is a spline 3^{rd}-order polynomial fit; the lateral lines represent the 95% confidence interval. Scatter points are averaged over intervals of 50 inhabitants. Bottom graphs: yearly RD coefficients. Vertical axis: point estimates of LLR regressions with optimal bandwidth computed following Calonico, Cattaneo, and Titiunik (2013a, 2013b). Horizontal axis: year. The central line is the point estimate; the lateral lines represent the 95% confidence interval.

Figure 3: Yearly RD estimates for policy outcomes and tax instruments



Notes. Yearly RD coefficients. Vertical axis: point estimates of LLR regressions with optimal bandwidth computed following Calonico, Cattaneo, and Titiunik (2013a, 2013b). Horizontal axis: year. The central line is the point estimate; the lateral lines represent the 95% confidence interval.

Appendix [For Online Publication Only]

This Appendix provides additional information and robustness checks, which are also discussed in the paper. In particular, we describe the characteristics and sources of the variables we use (Table A1), and we present further robustness checks:

- diff-in-disc estimates with covariates (Table A2);
- balance tests of time-invariant municipal characteristics (Table A3);
- diff-in-disc estimates on potentially endogenous variables (Table A4);
- falsification tests for the heterogeneity analysis (Table A5, Table A6, and Table A7);
- heterogeneity analysis for the main budget components (Table A8);
- test of the continuity of the density at 5,000 in the 1991 Census, in the 2001 Census, and with respect to the difference between the two Censuses (Figure A1);
- placebo tests based on permutation methods (Figure A2 and Figure A3).

Table A1: Variables' description and sources

Variable	Definition and measure	Available from-to	Source
Deficit	Expenditure minus revenues Per-resident; 2009 Euros	1997-2004	IMI Financial reports, authors' calculations
Fiscal gap	Expenditure minus revenues (net of central transfers and debt service) Per-resident; 2009 Euros	1998-2004	IMI Financial reports, authors' calculations
Current outlays	Total current expenditure Per-resident; 2009 Euros	1998-2004	IMI Financial reports, *Quadro 4*
Capital outlays	Total capital expenditure Per-resident; 2009 Euros	1998-2004	IMI Financial reports, *Quadro 5*
Debt service	Interest payments on outstanding debt Per-resident; 2009 Euros	1998-2004	IMI Financial reports, *Quadro 4*
Taxes	Total tax revenues Per-resident; 2009 Euros	1997-2004	IMI Financial reports, *Quadro 2*
Fees & tariffs	Total revenues from fees and tariffs Per-resident; 2009 Euros	1997-2004	IMI Financial reports, *Quadro 2*
Central transfers	Total transfers by the central state Per-resident; 2009 Euros	1997-2004	IMI Financial reports, *Quadro 2*
Other revenues	Residual category Per-resident; 2009 Euros	1997-2004	IMI Financial reports, authors' calculations
Real estate tax rate	The tax rate on real estate From 0.004 to 0.007 of the home value	1997-2004	IFEL-ANCI
Income tax surcharge	Municipal income tax surcharge Up to 0.6% of the taxable income	1999-2004	ME-DF

Notes: IMI stands for Italian Ministry of the Interior; IFEL-ANCI stands for Institute for the Local Finance and Economy of the National Italian Association of Municipalities; ME-DF stands for Italian Ministry of the Economy, Department of Finance.

Table A1: Variables' description and sources (cont'd)

Variable	Definition and Measure	Available from-to	Source
Census population	Census population of the municipality	1991 and 2001	ISTAT
Young cohorts	Ratio of residents aged 0–14 over resident population Fraction at municipality level	1998-2004	ISTAT
Speed of public good	Paid over committed current expenditures Fraction at municipality level	1999-2004	IMI Financial reports, authors' calculations
Area size	Municipal area size In km^2	1999-2004	IMI
Sea level	Municipal sea level In meters	1999-2004	IMI
Taxable income	Municipal taxable income mean Per-resident; 2009 Euros	1999-2004	ME-DF
Female Mayor	Equal to 1 if the mayor in office is a woman Dummy variable	1999-2004	IMI Register of local politicians
Mayor's age	Age of the mayor Number of years	1999-2004	IMI Register of local politicians
Mayor's schooling	Years of choosing of the mayor in office Number of years	1999-2004	IMI Register of local politicians
Mayor's tenure	Experience of the mayor in office Number of mandates	1999-2004	IMI Register of local politicians
Term limit	Equal to 1 if the mayor in office faces term limit Dummy variable	1999-2004	IMI Register of local politicians

Notes: IMI stands for Italian Ministry of the Interior; IFEL-ANCI stands for Institute for the Local Finance and Economy of the National Italian Association of Municipalities; ME-DF stands for Italian Ministry of the Economy, Department of Finance.

Table A2: The effect of relaxing fiscal rules, estimates with covariates

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
A. Fiscal discipline				
Deficit	16.220**	9.473**	21.094**	23.396*
	(7.656)	(4.140)	(9.283)	(12.579)
Obs	2,266	5,858	6,300	6,300
Fiscal gap	81.089***	48.296***	96.748***	118.962***
	(27.760)	(18.592)	(32.595)	(41.570)
Obs.	1,926	3,438	6,300	6,300
B. Expenditures				
Current outlays	-86.666	-9.246	-44.755	-68.705
	(57.987)	(28.486)	(50.871)	(72.117)
Obs.	1,656	4,112	6,300	6,300
Capital outlays	48.319	31.130	102.644	224.342*
	(87.377)	(75.384)	(101.934)	(136.181)
Obs.	1,500	3,974	6,300	6,300
Debt service	-2.674	-1.275	-1.962	-2.298
	(7.743)	(3.028)	(6.145)	(9.151)
Obs.	1,570	4,908	6,300	6,300
C. Revenues				
Taxes	-65.927**	-42.825**	-56.863**	-93.975***
	(27.474)	(18.377)	(24.024)	(30.751)
Obs.	1,350	2,810	6,300	6,300
Fees & tariffs	-9.613	-0.960	-2.992	-10.241
	(10.782)	(6.799)	(10.049)	(13.341)
Obs.	1,862	3,238	6,300	6,300
Central transfers	48.096**	33.136**	73.692**	93.268**
	(22.371)	(16.425)	(29.036)	(36.275)
Obs.	2,212	3,438	6,300	6,300
Other revenues	-30.357	-30.224	20.996	140.891
	(103.221)	(55.204)	(112.937)	(155.552)
Obs.	1,420	5,858	6,300	6,300
D. Tax instruments				
Real estate tax rate	-0.055**	-0.027*	-0.058**	-0.060*
	(0.026)	(0.016)	(0.026)	(0.033)
Obs.	1,636	3,806	6,300	6,300
Income tax surcharge	-0.035	-0.040	-0.051	-0.103**
	(0.043)	(0.026)	(0.042)	(0.052)
Obs.	1,206	2,594	4,588	4,588

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on policy outcomes and tax instruments below 5,000 after 2001. Specifications augmented by controlling for covariates: yearly dummies, macro areas dummies (i.e. North West, North East, South), area size (in km^2), and sea level (in meters). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. All policy outcomes are per capita and in 2009 Euros. Tax instruments are in percentage points. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A3: Balance tests of time-invariant characteristics

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
North-West	0.146	0.115	0.043	0.167
	(0.090)	(0.102)	(0.122)	(0.152)
Obs.	1,782	1,350	6,300	6,300
North-East	-0.065	0.009	0.042	-0.039
	(0.076)	(0.070)	(0.095)	(0.115)
Obs.	1,860	2,220	6,300	6,300
Center	-0.036	-0.074	-0.071	-0.181
	(0.088)	(0.056)	(0.106)	(0.133)
Obs.	1,710	5,430	6,300	6,300
South	-0.034	-0.094*	-0.013	0.053
	(0.076)	(0.048)	(0.088)	(0.110)
Obs.	1,878	4,494	6,300	6,300
Area size	4.323	3.261	0.043	-7.710
	(8.924)	(9.703)	(10.199)	(11.734)
Obs.	2,076	1,812	6,300	6,300
Sea level	33.969	5.542	-17.736	9.258
	(32.229)	(26.588)	(46.537)	(59.757)
Obs.	1,350	3,168	6,300	6,300

Notes. Municipalities between 3,500 and 7,000 inhabitants. Diff-in-disc estimates. All time-invariant characteristics are dummies except area size (in km^2) and sea level (in meters). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A4: Balance tests of potentially endogenous characteristics

	LLR CCT	LLR CV	Spline poly 3^{rd}	Spline poly 4^{th}
Taxable income	-743.050	-299.327	-230.188	-19.672
	(741.388)	(407.922)	(679.745)	(857.344)
Obs.	1,584	4,400	6,300	6,300
Female mayor	-0.027	-0.072	-0.112	-0.068
	(0.086)	(0.051)	(0.079)	(0.095)
Obs.	1,560	4,626	6,300	6,300
Mayor's age	-2.437	-0.126	-0.770	-0.907
	(3.173)	(1.697)	(3.114)	(3.930)
Obs.	1,678	4,634	6,300	6,300
Mayor's schooling	0.737	0.398	0.818	1.192
	(0.939)	(0.576)	(0.929)	(1.155)
Obs.	1,578	3,970	6,300	6,300
Mayor's tenure	1.376	0.454	0.444	1.287
	(1.510)	(0.742)	(1.497)	(1.848)
Obs.	1,736	5,858	6,300	6,300
No. of parties	-0.723	-0.131	-0.498	-0.958
	(0.516)	(0.390)	(0.485)	(0.621)
Obs.	1,470	2,656	6,300	6,300
Term limit	-0.084	-0.081	-0.136	-0.123
	(0.168)	(0.083)	(0.150)	(0.184)
Obs.	1,340	4,908	6,300	6,300
Young cohorts	-0.114	-0.132	-0.065	-0.083
	(0.124)	(0.180)	(0.154)	(0.198)
Obs.	2,554	1,302	6,300	6,300
Efficiency	0.131	0.123	0.137	0.161
	(0.162)	(0.092)	(0.156)	(0.196)
Obs.	1,626	4,394	6,300	6,300

Notes. Municipalities between 3,500 and 7,000 inhabitants. Diff-in-disc estimates. Taxable income at the municipal level is per capita and in 2009 Euros; mayor's age, schooling, and tenure are expressed in years; female mayor and term limit are dummies; number of parties refer to political parties seating in the city council. See the Appendix Table A1 for a precise description of all these variables. Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—CCT—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—CV—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A5: The political economy of deficit bias, part I – Falsification test

	LLR *CCT*	LLR *CV*	Spline poly 3^{rd}	Spline poly 4^{th}
		With two parties or less:		
Deficit	1.544	3.875	1.775	-9.176
	(7.090)	(9.182)	(13.148)	(15.106)
Obs.	1,959	1,136	1,959	1,959
		With more than two parties:		
Deficit	3.184	-3.329	-2.061	2.368
	(6.795)	(6.865)	(11.445)	(12.799)
Obs.	2,217	1,448	2,217	2,217
Difference between the two subsamples	1.640	-7.204	-3.836	11.544
Wald test p-value without covariates	0.625	0.166	0.244	0.166
		With binding term limit:		
Deficit	-12.302	-7.996	2.511	-24.188
	(10.465)	(12.240)	(17.210)	(22.310)
Obs.	1,211	701	1,211	1,211
		Without binding term limit:		
Deficit	5.146	-3.591	-2.964	5.388
	(6.806)	(8.215)	(13.713)	(16.833)
Obs.	2,965	1,807	2,965	2,965
Difference between the two subsamples	17.448	4.405	-5.475	29.576
Wald test p-value without covariates	0.175	0.872	0.806	0.297

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1997 and 2000. Diff-in-disc estimates of the (false) impact of introducing fiscal rules on policy outcomes below 5,000 after 1999 (when no discontinuity was introduced by the DSP) in different subsamples (that is, above vs. below median number of parties; binding vs. non-binding term limit). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—*CCT*—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—*CV*—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* is not available because of data limitations before 1999. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A6: The political economy of deficit bias, part II – Falsification test

	LLR *CCT*	LLR *CV*	Spline poly 3^{rd}	Spline poly 4^{th}
		Young cohorts above median:		
Deficit	4.136	-6.331	0.423	2.196
	(6.672)	(8.794)	(11.430)	(12.944)
Obs.	2,598	1,583	2,598	2,598
		Young cohorts below median:		
Deficit	-1.682	0.851	10.727	0.196
	(7.204)	(9.639)	(15.543)	(21.774)
Obs.	1,578	950	1,578	1,578
Difference between the two subsamples	-5.818	7.182	10.304	-2.000
Wald test p-value without covariates	0.564	0.608	0.607	0.940

Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1997 and 2000. Diff-in-disc estimates of the (false) impact of introducing fiscal rules on policy outcomes below 5,000 after 1999 (when no discontinuity was introduced by the DSP) in different subsamples (that is, above vs. below median percentage of young cohorts). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—*CCT*—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—*CV*—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* is not available because of data limitations before 1999. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A7: Fiscal restraints and budget management – Falsification test

	LLR *CCT*	LLR *CV*	Spline poly 3^{rd}	Spline poly 4^{th}
	Speed of public good provision above median:			
Deficit	3.500	-1.596	6.135	0.430
	(4.898)	(6.357)	(10.751)	(13.740)
Obs.	2,283	1,445	2,283	2,283
	Speed of public good provision below median:			
Deficit	2.812	-9.536	-5.696	-2.743
	(9.280)	(11.638)	(15.103)	(15.623)
Obs.	1,893	1,219	1,893	1,893
Difference between the two subsamples	-0.688	-7.940	-11.831	-3.173
Wald test p-value without covariates	0.947	0.469	0.501	0.871

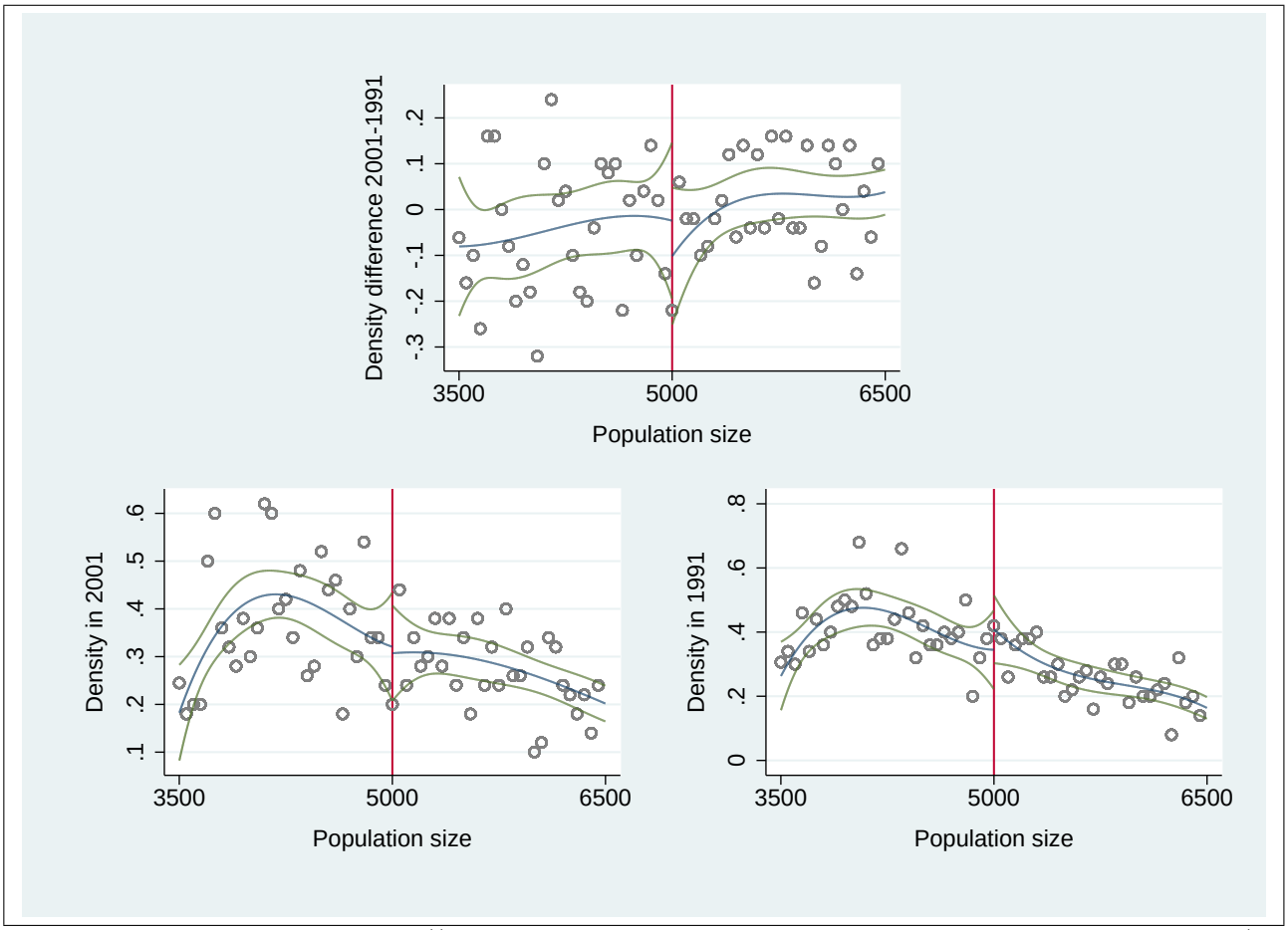
Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1997 and 2000. Diff-in-disc estimates of the (false) impact of introducing fiscal rules on policy outcomes below 5,000 after 1999 (when no discontinuity was introduced by the DSP) in different subsamples (that is, above vs. below median speed of public good provision). Estimation methods: Local Linear Regression (LLR) with two optimal bandwidth h, as in equation (5); spline polynomial approximation with 3^{rd}-order or 4^{th}-order polynomial, as in equation (6). The optimal bandwidth h is estimated either following Calonico, Cattaneo, and Titiunik (2013a, 2013b)—*CCT*—in the first column, or implementing the cross-validation algorithm proposed by Ludwig and Miller (2007)—*CV*—in the second column. The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* is not available because of data limitations before 1999. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A8: The composition of fiscal adjustment – Heterogeneity

	More than two parties	Binding term limit	Young cohorts	Speed of public good provision
Current Outlays				
Difference between the two subsamples	-16.414	150.782	-9.878	77.165
Wald test p-value without covariates	0.736	0.124	0.924	0.497
Wald test p-value with covariates	0.829	0.334	0.948	0.479
Capital outlays				
Difference between the two subsamples	-308.667	310.804	38.767	387.871
Wald test p-value without covariates	0.682	0.088	0.852	0.061
Wald test p-value with covariates	0.732	0.190	0.648	0.025
Taxes				
Difference between the two subsamples	-135.690	-1.828	-14.552	-90.703
Wald test p-value without covariates	0.032	0.970	0.773	0.068
Wald test p-value with covariates	0.024	0.857	0.102	0.229

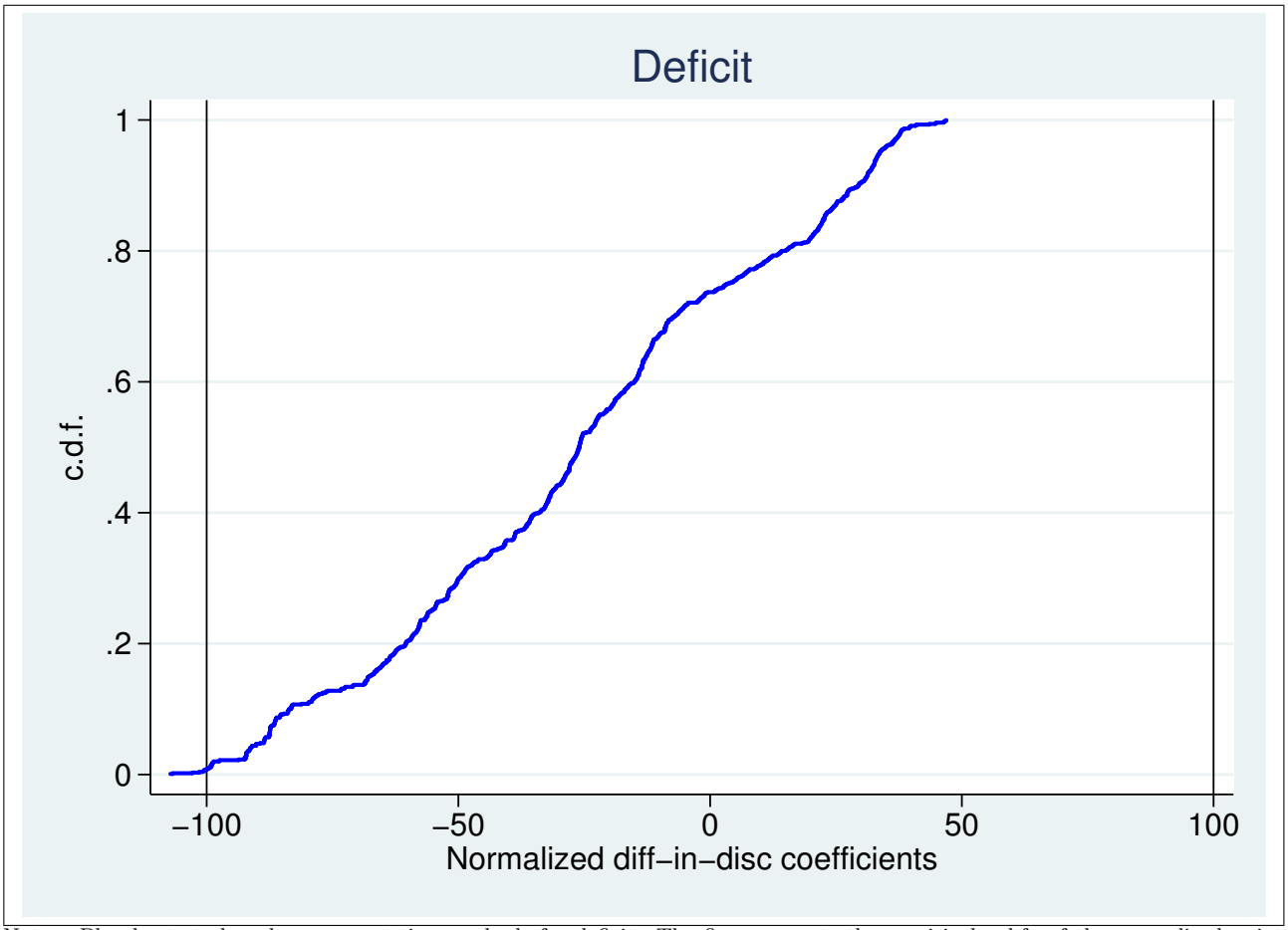
Notes. Municipalities between 3,500 and 7,000 inhabitants; budget years between 1999 and 2004. Diff-in-disc estimates of the impact of relaxing fiscal rules on policy outcomes below 5,000 after 2001 in different subsamples (as specified in each column heading). Estimation method: spline polynomial approximation with 3^{rd}-order polynomial, as in equation (6). The *Wald test p-value without covariates* evaluates whether the estimates are statistically different in the two subsamples. The *Wald test p-value with covariates* is not available because of data limitations before 1999. All variables are per capita and in 2009 Euros. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure A1: Density tests



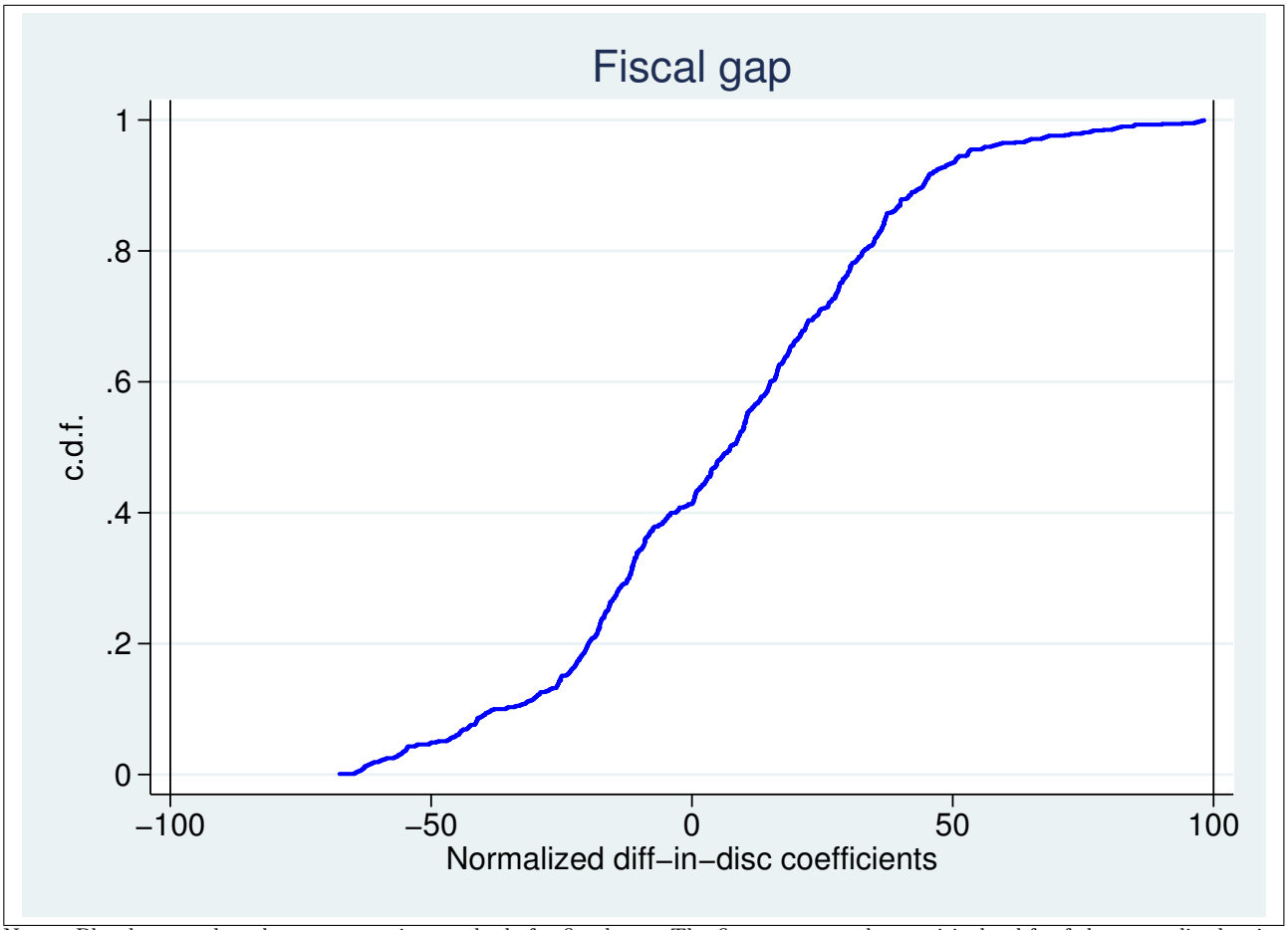
Notes. Test of the continuity at 5,000 of: (i) the difference between the density in the 2001 Census and in the 1991 Census (top graph); (ii) the density in the 2001 Census (bottom left graph); and (iii) the density in the 1991 Census (bottom right graph). The central line is a spline 3^{rd}-order polynomial fit in population size; the lateral lines represent the 95% confidence interval. Scatter points are averaged over intervals of 50 inhabitants.

Figure A2: Placebo tests for deficit



Notes. Placebo tests based on permutation methods for deficit. The figure reports the empirical c.d.f. of the normalized point estimates from a set of diff-in-disc estimations at 500 false thresholds below and 500 false thresholds above the true threshold at 5,000 (namely, any point from 4,900 to 4,400 and any point from 5,100 to 5,600). Estimation method: spline polynomial approximation with 3^{rd}-order polynomial. The vertical lines indicate our benchmark estimate for deficit from Table 4 (i.e., true coefficient normalized to 100) and its negative value.

Figure A3: Placebo tests for fiscal gap



Notes. Placebo tests based on permutation methods for fiscal gap. The figure reports the empirical c.d.f. of the normalized point estimates from a set of diff-in-disc estimations at 500 false thresholds below and 500 false thresholds above the true threshold at 5,000 (namely, any point from 4,900 to 4,400 and any point from 5,100 to 5,600). Estimation method: spline polynomial approximation with 3^{rd}-order polynomial. The vertical lines indicate the normalized benchmark estimate for fiscal gap from Table 4 (i.e., true coefficient normalized to 100) and its negative value.