A Congested Budget: The Fiscal Commons in Italian Municipalities*

Luca Bellodi† Massimo Morelli‡

Abstract

The literature on fiscal commons is unanimous on theory though discordant on findings. Theory predicts more politicians lead to inefficiently large programmes and hence overspending but findings are mixed. With novel data on Italian municipalities, we estimate the effect of the size of local councils and executive committees on a battery of planned and actual budget outcomes. We leverage a reform that introduced a new temporary population threshold where the size of local councils and executive committees changed discontinuously and estimate treatment effects with a difference-in-discontinuities design. We establish three novel results: *i*) more politicians lead to both more spending and more revenues, leaving deficit unchanged; *ii*) the effects disappear when looking at what politicians actually spend and collect; and *iii*) the difference between actual and planned budget is smaller in municipalities with a larger share of graduate politicians and bureaucrats, suggesting capacity deficits prevent politicians from implementing the planned budget.

*We are thankful to Tom O'Grady, Lucy Barnes, Adam Harris, Tim Hicks, Alexander Coppock, Rafael Ch, Gianmarco Daniele, participants at the UCL Graduate Students Seminar, and participants at MPSA Conference 2021. We also wish to thank Monica Mantovani, Daniele Formiconi (ANCI), and Cristiano Durantini (Ministry of the Interior) for helpful guidance in surveying legislation and accessing data.

†Bocconi University, DONDENA Center, and UCL Political Science.

‡Bocconi University, IGIER, CEPR, PERICLES.

In the last decades, various reforms have been proposed aimed at reducing the number of elected politicians in the attempt to restore trust in government or make the legislative process more efficient. Scholarly work on the economic consequences of government composition generally follows the theoretical gist of the fiscal commons problem or the law of 1/N (Weingast, Shepsle, and Johnsen 1981), which posits that – because of the generalised costs (through taxation) and the geographically concentrated benefits of legislative policy-making (the constituencies which benefit from the policy) – more electoral districts/representatives lead to inefficiently large projects, and hence overspending. In this paper, we use data on Italian municipalities and show that more politicians lead to a congested budget, with larger *planned* expenditures and revenues that however do not translate into larger deficits and are not implemented in practice.

Weingast, Shepsle, and Johnsen (1981) present a model of distributive politics where elected politicians care about the public projects that benefit their districts. By assumption, these projects are all passed and funded through general taxation. The costs of the projects are equally shared among the districts, whereas the benefits are geographically concentrated. As a result, elected politicians will demand inefficiently large projects, with inefficiency increasing in the number of representatives. This theory inspired a large empirical scholarship covering different countries and levels of government, but the findings are nonetheless mixed (for a review and meta analysis of the studies, see Freire et al. 2021). We believe that important sources of heterogeneity are hidden behind the seemingly contrasting results documented in the literature. In this paper we focus on two of them. First, we estimate the effect of the size of municipal government bodies on expenditures, revenues, and deficit, hence providing a comprehensive picture of fiscal policy. Second, we distinguish between what the government plans to spend at the be-

ginning of the fiscal year and what government actually spends at the end of the fiscal year, allowing for spending and collection capacity deficits to arise. We believe that these two additional layers can explain why some studies find support for the theory and why others fail to do so.

Empirically, we study the context of Italian municipalities, where the size of government bodies – as well as many other policies – changes discontinuously at various population thresholds. For identification, we exploit a natural experiment offered by a reform implemented in the summer of 2011 which temporarily introduced a new population threshold at 5,000 inhabitants, which allows us to compare changes in budget outcomes at the cutoff with and without the temporary threshold in a difference-in-discontinuities design.

We find that treated municipalities in the reform period (with about 4 politicians more) *plan* to spend more, with expenditures increasing by almost 600 euros per capita. The increase in spending is coupled by an equal-sized increase in revenues, leaving deficit unchanged. This is the first contribution of our paper: larger expenditures do not necessarily mean overspending. However, the effects disappear when looking at *actual* budget outcomes. We document a congestion effect, whereby more politicians pass larger budgets that cannot be implemented. In the data, we find support for a capacity deficit mechanism. The spending and collection capacity of municipalities with larger shares of graduate politicians and bureaucrats is between 40 and 20 percentage points larger compared to municipalities with a lower share of graduate politicians and bureaucrats. The congestion effect we document can therefore be compensated by high-capacity governments. These findings introduce important scope conditions that should be considered when interpreting existing and future tests of the law of 1/N.

Institutional Context and Data

The organisation of municipal governments in Italy consists of a directly elected mayor – who appoints an executive committee – and a directly elected local council with quasi-legislative prerogatives and political oversight functions. Municipal elections are held every five years. Mayoral candidates run for office supported by one or more party-lists of candidates to the local councils. Voters can express a preference for candidates to the local council, and – according to the seats awarded to the respective party-list – the candidates with the largest number of votes are elected. The law defines the size of the local council and a cap to that of executive committees as a function of census population.

Municipal governments are responsible for a whole range of services, from municipal police to housing, schooling, welfare politics, waste management and infrastructure development, and every decision with financial implications must have financial coverage as specified in the budget, which lists the type and source of available resources and how the government intends to spend them. Expenditures are financed by municipal revenues, which consist of financial transfers from regional and central government, local taxes and tariffs, as well as other economic activities of the municipality. The budget cycle starts with the executive committee presenting a budget proposal at the beginning of every fiscal year, in which it outlines the revenues and expenditures *planned* for the next three years (i.e, planned budget). At the end of the fiscal year, municipal governments approve a final budget which accounts for the *actual* expenditures and revenues incurred by the government throughout the year (i.e., actual budget). Importantly, the budget must be approved by the local council, which in turn has a significant influence on the fiscal decisions of the municipality.

We assembled a dataset consisting of budget data – both planned and actual – and

government composition for all Italian municipalities from 1998 to 2015. We web-scraped budget data from the repositories of the Local Public Finance Directorate of the Ministry of the Interior and focus on three key planned and actual outcomes: expenditures, revenues, and deficit, measured as the difference between expenditures and revenues. Data on government composition (number of councillors and members of the executive committee) comes form the Local Administrators Database of the Ministry of the Interior.[1]

Research Design

The size of government bodies changes discontinuously at seven census population thresholds: 3, 10, 30, 100, 250, 500, 1,000 thousand inhabitants. A survey of the legislation reveals that the wage of the mayor, members of the executive committee, and of local councillors jump discontinuously at the same population thresholds (see Section A.3 in the Online Appendix). As a result, comparing units just above and below these thresholds with a cross-sectional RD estimator would not allow to distinguish the effect of more politicians from that of better paid politicians. However, in the summer 2011 the Berlusconi IV government passed a law that introduced a new threshold at 5,000 inhabitants at which the composition of government bodies was supposed to change, which remained into force until 2014. The change brought about by the reform is displayed in Figure 1.

Wage policies and – until 2012 – a fiscal rule aimed at controlling budgetary balances change at the same cutoff, but the temporary difference in the size of government bodies at the same threshold allows us to compare changes in fiscal outcomes at the cutoff before and after the reform, thus isolating the effect of the number of politicians from the confounding policies that jump simultaneously at the same cutoff. We subset the data to

[1]Description of the sample of municipalities, data sources, and descriptive statistics are reported in the Online Appendix (see Section A.1 and Section A.2).

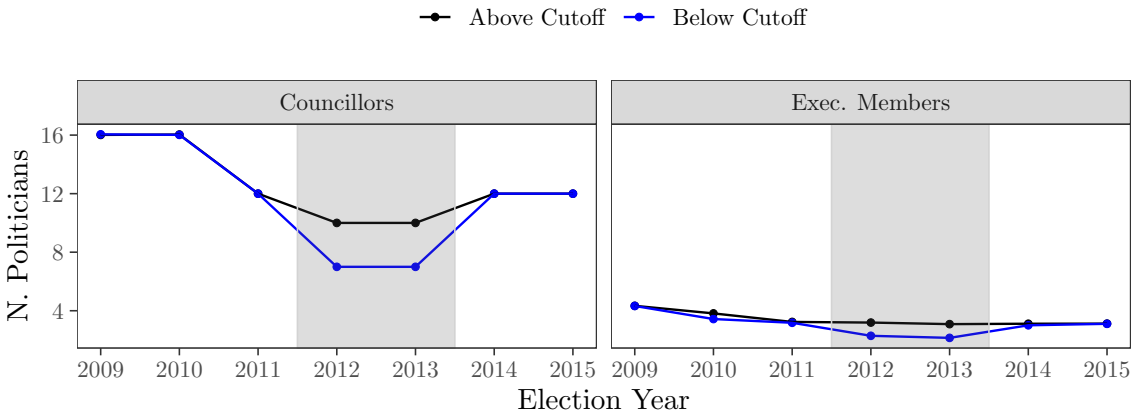

Figure 1: Average number of local councillors and members of executive committees for municipalities in the 3-10,000 population band before and after the reform. The decrease in 2011 is the result of a different reform which affected all municipalities without establishing a new population threshold.

the post-fiscal rule period when only wage policies confound the effect of the number of politicians (fiscal years 2013-2015). We then estimate the effect of having more politicians with and without the reform in a difference-in-discontinuities design (Grembi, Nannicini, and Troiano 2016; Eggers et al. 2018).

The Difference-in-Discontinuities Estimator

Under the reform, in $T_t = 1$, the RD estimator identifies the effect of both wage policies W_i and number of politicians P_i on the outcome Y_i. The compound estimand is given by $\alpha_{RD(c)} = \lim_{\epsilon \downarrow c} \mathbb{E}[Y_i | X_i = \epsilon, T = 1] - \lim_{\epsilon \uparrow c} \mathbb{E}[Y_i | X_i = -\epsilon, T = 1]$. Grembi, Nannicini, and Troiano (2016) show that, under additional local assumptions, information on the periods without the compound treatment ($T_t = \{0, 2\}$) allows to isolate the effect of P_i from that of W_i. Let δ_{RD} be the causal effect of W_i when $T_t \in \{0, 2\}$. To identify the causal effect of P_i, we combine both the discontinuous variation at $C_i = 5,000$ and the time variation when T moves from 0 to 1 and then to 2: $\tau_{DD(c)} = \alpha_{RD(c)} - \delta_{RD(c)}$. $\tau_{DD(c)}$ is the difference-in-discontinuities estimator (hereafter, diff-in-disc). Importantly, the target estimand is narrower than the standard RD estimand, for it is conditional on

the realisation of the confounding treatment. $\tau_{DD(c)}$ is thus the effect at the cutoff of having more politicians for municipalities that also have better paid politicians.

Two additional assumptions compared to the two continuity assumptions of RD designs must hold for the diff-in-disc to be unbiased. The first one is that the effect of W_i and P_i is constant over time. In other words, it requires that units just above and below the threshold would have held a parallel trend had P_i not been introduced. The second assumption is that the effect of P_i does not depend on the confounding policy W_i. One way in which this assumption would be violated is if politicians above and below the cutoff, who are paid differently, reacted to a change in the number of politicians differently. These assumptions are indirectly tested and discussed in Section A.8 of the Online Appendix, whereas support for the continuity of density and potential outcomes assumptions are reported in Section A.7.

We estimate $\tau_{DD(c)}$ with local polynomial methods, fitting linear WLS regressions separately on the observations above and below the cutoff and before, during, and after the reform (in $T = 0, 1, 2$ separately). Weights are determined by the triangular kernel function based on the ratio between the distance of unit i from the cutoff C_i and the mean-squared-error minimising bandwidth h (i. e., $w_i = 1 - \frac{X_{it} - C_i}{h}$) (Cattaneo, Idrobo, and Titiunik 2019).[2] The RD estimates equal the difference in the intercepts at the cutoff in every time period. Subsequently we estimate the difference in the estimated effects across the three different periods.

Results

The results are displayed in Figure 2, which shows in the top panels the RD estimates for each period (before, during, and after the reform) and in the bottom panel the diff-in-disc

[2]Results are robust to alternative bandwidths, see Figure A.2 in the Online Appendix.

estimates, estimated as the difference in the effects between the reform and pre-reform samples (black coefficients) and the post-reform sample (blue coefficients).[3] Consistently with the theoretical predictions, for municipalities above the 5,000 cutoff, an addition of three councillors (from 7 to 10) and the possibility to appoint one additional member of the executive committees (from 3 to 4) has large effects on planned fiscal outcomes. Expenditures per capita are larger by 596 and 586 euros in the reform period compared to the pre and post reform, respectively. Revenues increase too and to a very similar extent (i.e., 583 and 608 euros per capita compared to the pre and post-reform period).[4] Importantly, the increased spending coupled by an equal-size increase in revenues leave planned deficit unchanged. Having approximately four politicians more does not lead to overspending or inefficiently large expenditures, for larger planned expenditures do not exceed larger planned revenues.

However, municipal governments might fail to realise the planned budget passed at the beginning of the fiscal year. When looking at the right-hand side panels, municipalities above the cutoff in the reform period have very similar budgets to municipalities above the cutoff before and after the reform. The difference between the estimates during and before/after the reform is much smaller compared to the planned budget, with standard errors being more than twice as large as the point estimates. More politicians lead to a congested budget, with larger spending and revenues that nonetheless do not materialise.[5]

[3]Full regression tables are reported in Section A.5 in the Online Appendix.

[4]In Figure **??** in the Online Appendix we show that the increased revenues are mostly driven by larger transfers from higher levels of government and not by increased taxation.

[5]In Section A.6 in the Online Appendix we show results are robust to excluding covariates, using alternative data from the National Institute of Statistics, accounting for the government term business cycle, and limiting the analysis to municipalities which held elections before 2013, for from December of 2012 gender quotas on candidate lists started to operate based on the same population threshold and might confound the effect of the number of politicians.

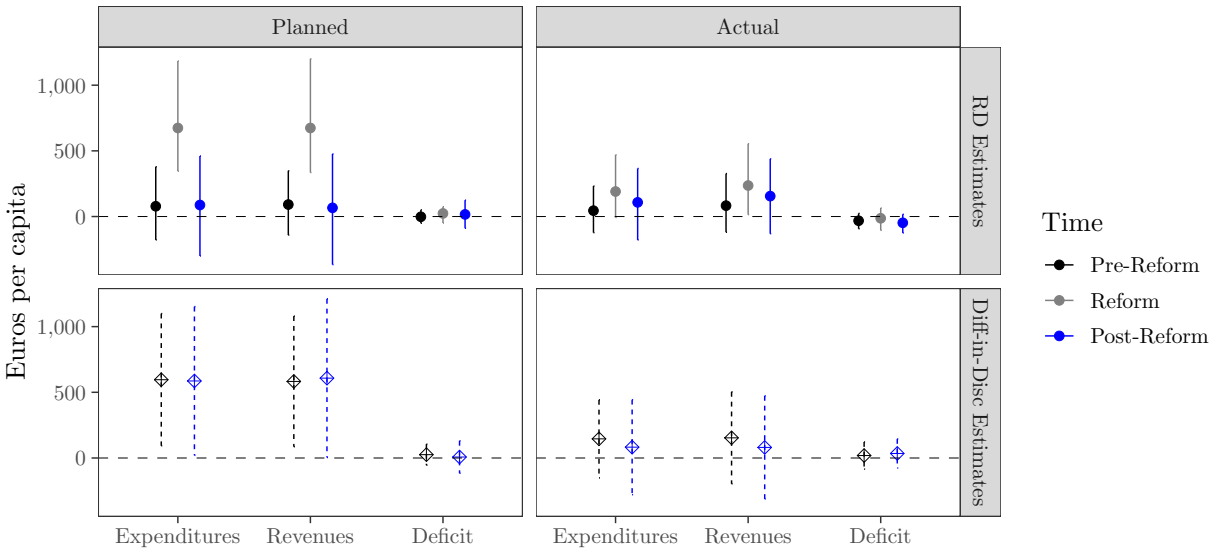

Figure 2: RD (top panels) and Diff-in-Disc (bottom panels) estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence intervals constructed using bias-correction with robust standard errors. Covariates include: log population density, log surface (sq.km), log surface at low, medium, and high hydro-geological risk (sq.km), gender and degree of mayor (dummy), northern-region indicator (dummy), year dummies. Analysis implemented with the *rdrobust* package in R.

Mechanism

The congested budget is characterised by larger expenditures and revenues that however to not materialise in practice, reducing the spending and collection capacity of government (measured as the ratio of actual and planned budget items). One key question is what is preventing them from implementing the planned budget. Collection and spending capacity deficit might be the product of politicians' inability to "get things done" or the resistance they face from under-performing bureaucracies. To test whether there is empirical support for this mechanism, we replicate the analysis on two separate samples of municipalities whose levels of politicians and bureaucratic capacity are above or below the median in the sample.

We proxy politicians' and bureaucratic capacity with the percentage of politicians and bureaucrats with a university degree and the number of bureaucrats. Figure 3 shows the diff-in-disc estimates in the two samples of municipalities above and below the median

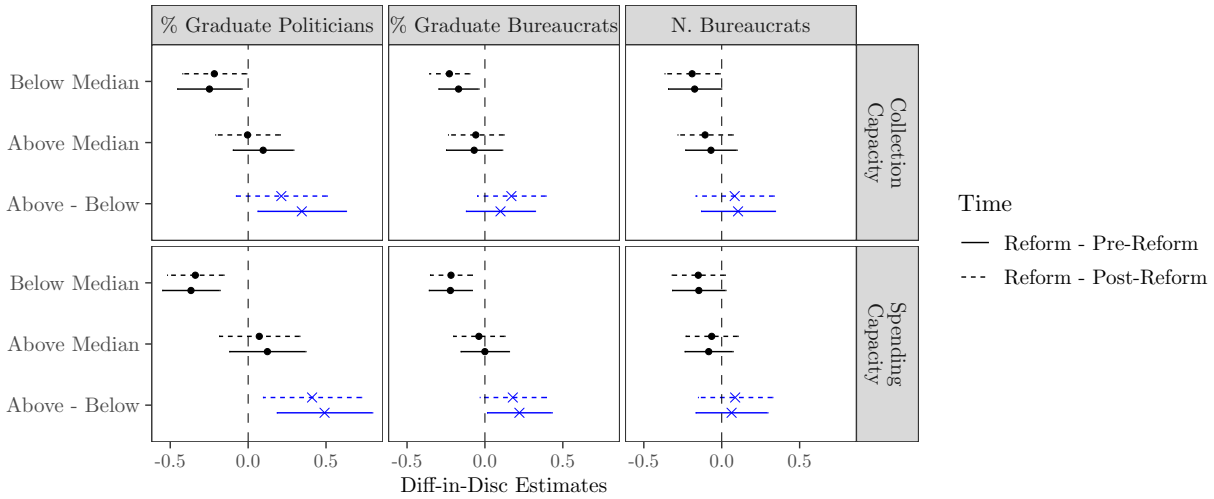

Figure 3: Diff-in-Disc estimates across two samples (units below and above median value of capacity covariates reported in panel labels) and difference in the estimated effects (blue coefficients). Outcomes are collection capacity (actual/planned revenues per capita) and spending capacity (actual/planned expenditures per capita) Same covariates and estimation as reported in Figure 2.

(black coefficients), as well as the difference between the two samples (blue coefficients) for each of the three capacity variables (top panels' labels). These tests show that spending and collection capacity sharply increase in municipalities with a larger share of graduate politicians by more than 40 and 20 percentage points, respectively. There is also little evidence that spending and collection capacity are larger in municipalities with a larger share of graduate bureaucrats, although the difference is distinguishable form zero at 95% level only for spending capacity in the reform - pre-reform period. Conversely, the size of the bureaucracy does not alter spending and collection capacity, suggesting capacity deficits are more driven by the quality rather than the quantity of government officials.

Discussion

In this article we contribute to the empirical literature on the fiscal consequences of government compositions, bringing two new perspectives: larger spending does not necessarily mean overspending and might not be implemented in practice. What is clear from the results is that more politicians lead to a congested budget. While planning to

expand their budgets, they fail to do so in practice. There is evidence this congestion is driven by a capacity deficit which prevents legislators from implementing the planned budget. When the share of politicians and bureaucrats with a university and degree is high, collection and spending capacity drastically increase. These findings should guide future empirical research aimed at testing the law of 1/N. We suggest researchers do not limit their analysis to expenditures and estimate the effects on both planned and actual budget. Finally, when researchers find null results, they should test whether the null masks heterogeneity across different levels of government capacity.

References

Cattaneo, Matias D, Nicolas Idrobo, and Rocío Titiunik. 2019. *A Practical Introduction to Regression Discontinuity Designs: Foundations.* Vol. I. Cambridge University Press.

Cattaneo, Matias D, Michael Jansson, and Xinwei Ma. 2020. "Simple Local Polynomial Density Estimators." *Journal of the American Statistical Association* 115 (531): 1449–55. https://doi.org/10.1080/01621459.2019.1635480.

Eggers, Andrew C., Ronny Freier, Veronica Grembi, and Tommaso Nannicini. 2018. "Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions." *American Journal of Political Science* 62 (1): 210–29. https://doi.org/10.1111/ajps.12332.

Freire, Danilo, Umberto Guarnier Mignozzetti, Catarina Roman, and Huzeyfe Alptekin. 2021. "The Effect of Legislature Size on Public Spending: A Meta-Analysis," 1–20. https://doi.org/10.31235/osf.io/xf7wp.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. "Do fiscal rules matter?" *American Economic Journal: Applied Economics* 8 (3): 1–30. https://doi.org/10.1257/app.20150076.

Weingast, Barry R, Kenneth A Shepsle, and Christopher Johnsen. 1981. "The Political Economy of Benefits and Costs: A Neoclassical Approach to Distributive Politics." *Journal of Political Economy* 89 (4): 642–64.

Appendix

Contents

A.1 Data Sources

We combine four sources of data. First, I obtained full lists of municipalities and unique codes directly from the Local Public Finance Directorate of the Ministry of Interior. I then used the unique codes to scrape budget data from the on-line ministerial repository. Second, I merged budget data with data on local government composition obtained from the Database on Local and Regional Administrators, curated by the Ministry of the Interior. Third, I obtained demographic data from the National Institute of Statistics, and fourth, data on the number and level of education of municipal employees from the national account of the Italian General Accounting Office. All datasets available at finanzalocale.interno.gov.it (budget data), dait.interno.gov.it/elezioni/open-data (politicians data), dat.istat.it (demographic data), and contoannuale.mef.gov.it (municipal employees).

A.2 Sample of Municipalities and Inflation Adjustment of Budget Data

From the total sample of municipalities, we removed the municipalities of the five special statute regions (Sicilia, Sardegna, Valle d'Aosta, Friuli-Venezia Giulia, and Trentino-Alto Adige/Südtirol) for the thresholds mechanism apply to those regions only insofar as they are compatible with their own special statutes, and because these municipalities are subject to financial constraints and rules that differ from those in force for the the remaining 15 regions.

All the outcome variables are per capita and adjusted to 2015 inflation level. Figures are therefore expressed in real terms and are comparable over time. I applied the consumer price index for currency evaluations as per January of every year (Jan. 2015 index = 100). Time-series downloaded from the archive of the National Institute of Statistics, www.istat.it/it/archivio/30440.

A.3 Policies Based on Population Thresholds: Legal Sources

Rule/Policy	**Time-Frame**	**Legal Source**	**Cutoffs [1,000]**
Wage Mayor	1998-2000 2001-Present	Art. 3 and Table A, Law 861/1985 Art. 1 and Table A, Min. Dec. 119/2000	3, 5, 10, 30, 50 1, 3, 5, 10, 30, 50
Wage Vice Mayor	2001-Present	Art. 3 and Table A, Min. Dec. 119/2000	1, 3, 5, 10, 30, 50
Wage Ex. Comm.	2001-Present	Art. 4, Min. Dec. 119/2000	1, 5, 50
Reimb. Councillors	1998-2000 2001-Present	Art. 10, Law 861/1985 Art. 1 and Table A, Min. Dec. 119/2000	30 1, 10, 30
Wage Pres. Council	2001-Present	Art. 5, Min. Dec. 119/2000	1, 15
Neigh. Councils	1998-2000 2001-Present	Art. 13, Law 142/90 Art. 17, Leg. Dec. 267/2000	30 (optional)
N. Financ. Auditors	1998-2000 2001-2006 2007-Present	Art. 57, Law 142/90 Art. 234, Leg. Dec. 267/2000 Art. 1(732) Law 296/2006	5 5 15
Director General	1998-2000 2001-2003	Art. 51-bis, Law 142/90 Art. 108, Leg. Dec. 267/2000	15
Fiscal Rule	2001-2012	Art. 53, Law 388/2000 and other subsequent laws	5
Gender Quotas	2013-Present	Art. 2, Law 215/2012	5, 15
Cut to Size of Government Bodies	2011-2014	Art. 2(184) Law 191/2009, Art. 1(2) Law 42/2010	1, 3, 5, 10, 30, 50
Introduction of 5,000 threshold	2011-2014	Art. 16(17) Decree-Law 138/2011	1, 3, 5, 10
Election Date	1999-Present	Art. 1 and 2 Law 182/1992	

Table A.1: Legal sources of all rules/polices based on population thresholds with reported cutoffs up to 50,000 inhabitants.

	Population Thresholds						
	1,000	3,000	5,000	10,000	15,000	20,000	30,000
Wage Policies							
Mayor	2001-Present	**X**	**X**	**X**			**X**
Vice Mayor	2001-Present		**X**	**X**			
Reimb. Councillor	2001-Present			**X**			
Exec. Comm.	2001-Present		**X**				
President of L Counc.	2001-Present				**X**		
Other Policies							
Neighbour. Councils							?
Health Centre						?	
Fiscal Rules			2001-12				
Gender Quotas			2013-Present		2013-Present		
Financial Auditors			1998-06		**X**		
Electoral System					**X**		
Director General					?		
Treatment							
Local Council Size	2011-13	**X**	2011-13	**X**			**X**
Executive Comm.	2011-13	**X**	2011-13	**X**			**X**

Table A.2: Rule/Policies that change with population thresholds with respective time-frame (entire time-frame when omitted) Thresholds larger than 30,000 omitted. **X** signifies deterministic change, ? signifies possible change, namely cases in which municipalities *may* adopt the rule or implement the policy if they want to. Hospital means whether the municipality can have a hospital or health centre and neighbourhood councils are local councils that can be established in any neighbourhood. Two important exceptions: wage/reimbursement thresholds at 1,000 introduced only from 2001.

A.4 Descriptive Statistics

Table A.3 reports descriptive statistics for the main variables for the total sample and the sample of municipalities in the 3-10,000 inhabitants population band.

	Full Dataset			**3-10,000 Population Band**		
N. Municipalities		8,451			2,083	
N. Observations		143,406			5,832	
Budget Item	**Mean**	**Median**	**SD**	**Mean**	**Median**	**SD**
Local Councillors	14.2	12.0	4.8	11.6	12.0	1.3
Members of Exec. Comm.	3.2	3.0	1.7	2.3	3.0	1.0
Planned Budget						
Expenditures pc	1,768.5	1,320.7	2,180.4	1,689.2	1,322.7	1,963.6
Revenues pc	1,753.6	1,306.2	1,972.7	1,678.0	1,306.0	1,958.6
Deficit pc	12.2	5.6	251.1	11.2	5.9	180.5
Actual Budget						
Expenditures pc	999.8	809.3	977.0	962.3	797.8	1,266.3
Revenues pc	1,035.9	825.2	1,103.4	982.4	803.3	1,313.7
Deficit pc	-36.1	-13.4	454.1	-20.2	-4.5	296.7

Table A.3: Descriptive statistics of main variables in the entire dataset and for the sample of units with census population between 3,001 and 10,000 inhabitants. Descriptive statistics of main variables. Deficit per capita measures are equal to the difference between total expenditures and total revenues divided by the resident population.

A.5 Regression Tables

	Planned Budget			Actual Budget		
Time	Expend.	Reven.	Deficit	Expend.	Reven.	Deficit
Pre-Reform						
Estimate	78.3	91.3	-1.9	44.8	82.9	-32.0
SE	(142.1)	(124.9)	(26.1)	(90.6)	(113.9)	(29.9)
p.value	0.480	0.407	0.996	0.546	0.363	0.247
h	758.1	967.6	517.1	1050.2	914.0	667.7
Obs. Used	506	717	352	764	676	453
Reform						
Estimate	674.2	674.2	23.5	190.4	236.2	-13.3
SE	(213.8)	(220.8)	(31.7)	(121.6)	(137.7)	(43.5)
p.value	0.000	0.001	0.688	0.057	0.039	0.645
h	535.8	515.1	552.4	780.9	661.9	585.1
Obs. Used	201	198	216	327	270	231
Post-Reform						
Estimate	87.9	66.3	16.1	108.0	155.6	-47.7
SE	(194.1)	(214.9)	(54.5)	(138.8)	(145.4)	(36.2)
p.value	0.679	0.797	0.746	0.499	0.289	0.136
h	603.1	635.8	756.8	653.4	675.6	722.6
Obs. Used	461	475	584	493	516	555

Table A.4: RD estimates as displayed in Figure 2 for each time period and each outcome separately. Estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth (i.e., h). Robust p.values computed using bias-correction with robust standard errors. Covariates include: log population density, log surface (sq.km), log surface at low, medium, and high hydro-geological risk (sq.km), gender and degree of mayor (dummy), northern-region indicator (dummy), year dummies. Analysis implemented with the *rdrobust* package in R.

	Reform - Pre-Rerom			Reform - Post-Reform		
	Expend.	Reven.	Deficit	Expend.	Reven.	Deficit
Planned Budget						
Diff.	595.9	582.8	25.4	586.3	607.8	7.3
SE	(256.8)	(253.7)	(41.0)	(288.8)	(308.1)	(63.1)
p.value	0.020	0.022	0.536	0.042	0.049	0.907
Actual Budget						
Diff.	82.4	80.6	34.4	82.4	80.6	34.4
SE	(184.5)	(200.2)	(56.6)	(184.5)	(200.2)	(56.6)
p.value	0.655	0.687	0.543	0.655	0.687	0.543

Table A.5: Difference in the RD estimates as displayed in Figure 2 for every outcome.

A.6 Robustness Tests

	Reform - Pre-Rerom			Reform - Post-Reform		
	Expend.	Reven.	Deficit	Expend.	Reven.	Deficit
Planned Budget						
Difference	607.0	619.3	-7.1	656.2	675.0	-13.6
SE	(304.8)	(317.3)	(38.7)	(330.4)	(357.4)	(63.7)
p.value	0.046	0.051	0.854	0.047	0.059	0.831
Actual Budget						
Difference	228.0	181.7	51.4	228.0	181.7	51.4
SE	(210.2)	(218.4)	(59.8)	(210.2)	(218.4)	(59.8)
p.value	0.278	0.405	0.390	0.278	0.405	0.390

Table A.6: Diff-in-Disc estimates without including covariates in RD estimation across each time period.

	Reform - Pre-Rerom			Reform - Post-Reform		
	Spending Capacity	Collection Capacity	Deficit	Spending Capacity	Collection Capacity	Deficit
Difference	-0.15	-0.10	-0.10	-0.17	-0.16	-0.38
SE	(0.05)	(0.06)	(0.11)	(0.06)	(0.06)	(0.32)
p.value	0.006	0.070	0.356	0.003	0.005	0.236

Table A.7: Diff-in-Disc estimates using alternative data for outcomes from the National Institute of Statistics. Spending capacity is the ratio between actual and planned expenditures; collection capacity is ratio between actual and planned revenues; deficit is administration remainder divided by planned revenues.

Because the analysis is performed on a sample covering three fiscal years, pre- and post-reform differences might be confounded by the year-of-term effect, with municipalities further down the government term more likely to spend more. However, we show that the estimates after removing the effect of the government term cycle on the outcomes are even larger.

	Reform - Pre-Rerom			**Reform - Post-Reform**		
	Expend.	Reven.	Deficit	Expend.	Reven.	Deficit
Planned Budget						
Difference	603.0	586.9	25.4	595.4	612.0	7.3
SE	(259.3)	(254.9)	(41.0)	(290.3)	(308.5)	(63.1)
p.value	0.020	0.021	0.536	0.040	0.047	0.907
Actual Budget						
Difference	87.0	82.8	34.4	87.0	82.8	34.4
SE	(185.1)	(200.5)	(56.6)	(185.1)	(200.5)	(56.6)
p.value	0.639	0.680	0.543	0.639	0.680	0.543

Table A.8: Diff-in-Disc estimates. Outcomes from RD estimates are the residuals of linear regressions of each outcome on the year-of-term variable. Same RD estimation reported in Table A.4.

Because from 2013 gender quotas on candidate lists started to operate based on the same population threshold, the difference between the reform and pre-reform periods might be confounded by the presence of more female councillors in the government. However, looking at the difference between the reform and post-reform estimates periods differences out the effect of gender quotas, for it remains in 2014 too, when the size of government bodies ceases to change at the 5,000 threshold. Moreover, the similar effect size of the diff-in-disc estimates in the pre- and post-reform period might suggest that the gender composition of the local council has a negligible effect on local public finance.

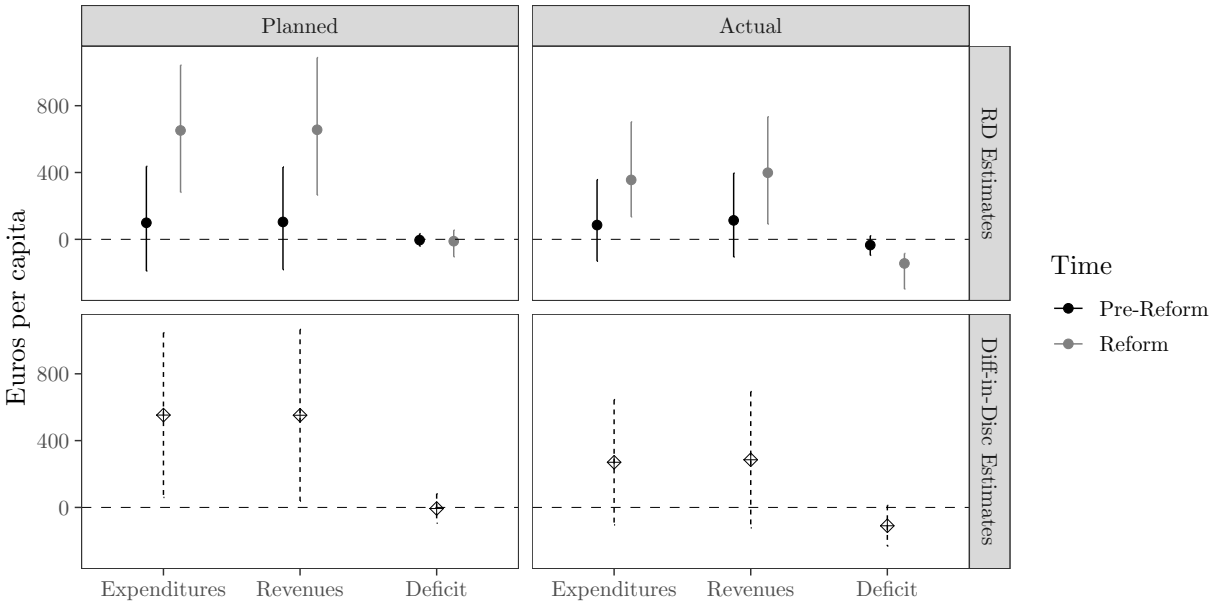

Figure A.1: RD (top panels) and Diff-in-Disc (bottom panels) estimates from sample of municipalities which held elections earlier than 26 December 2012, when the gender quotas entered into force for municipalities above the 5,000 population threshold. Estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence intervals constructed using bias-correction with robust standard errors. Covariates include: log population density, log surface (sq.km), log surface at low, medium, and high hydro-geological risk (sq.km), gender and degree of mayor (dummy), northern-region indicator (dummy), year dummies. Analysis implemented with the *rdrobust* package in R.

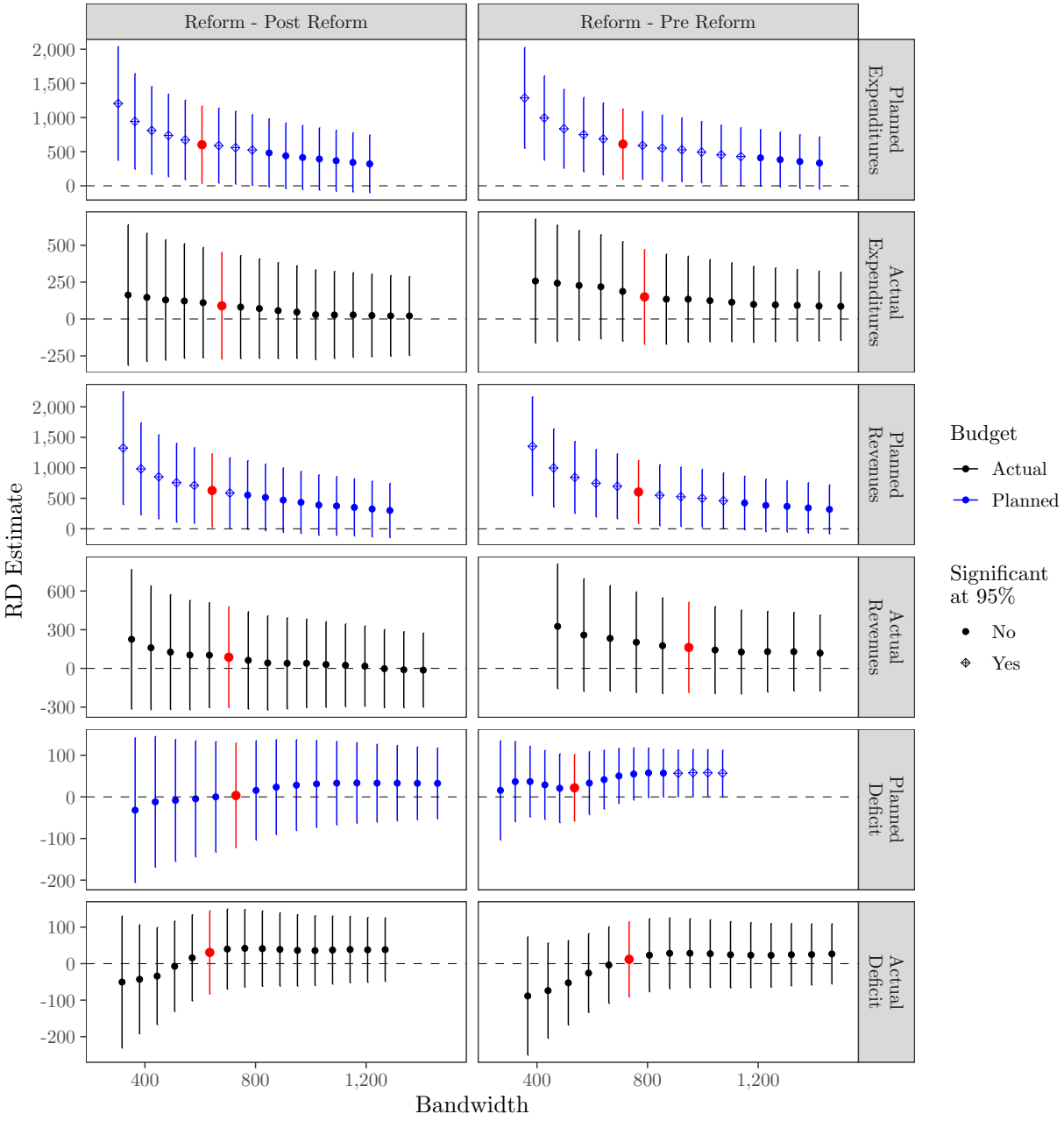

Figure A.2: Diff-in-Disc estimates with 95% CI estimated with alternative bandwidths. Red coefficients estimated with MSE-optimal bandwidth. Underlying RD estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence interval constructed using bias-correction with robust standard errors. Same covariates included in Table A.4.

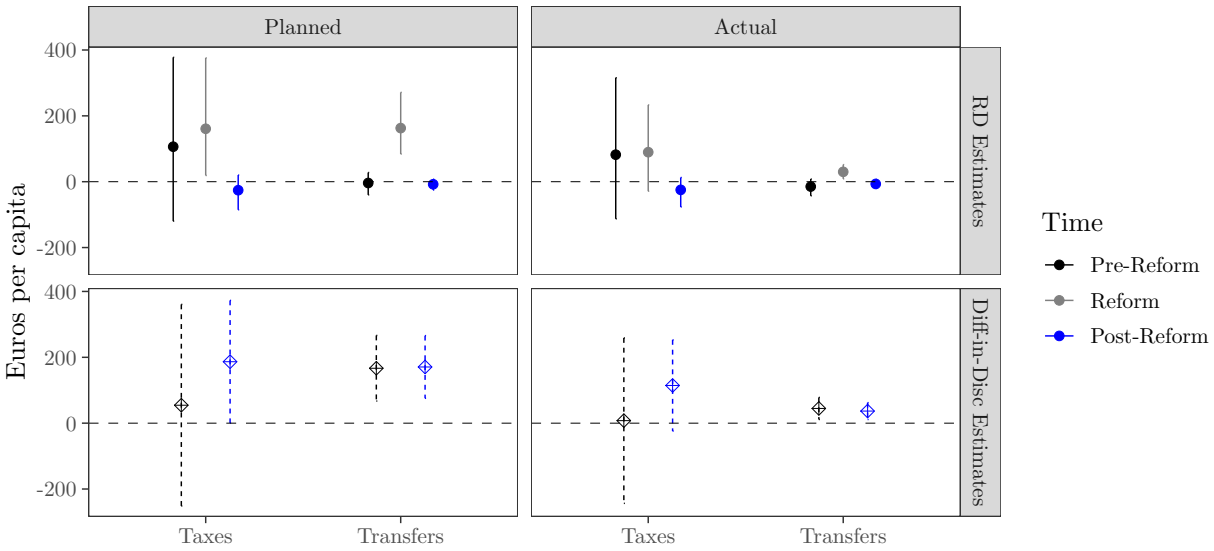

Figure A.3: RD (top panels) and Diff-in-Disc (bottom panels) estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence intervals constructed using bias-correction with robust standard errors. Covariates include: log population density, log surface (sq.km), log surface at low, medium, and high hydro-geological risk (sq.km), gender and degree of mayor (dummy), northern-region indicator (dummy), year dummies. Analysis implemented with the *rdrobust* package in R.

A.7 Validity of RD Estimator

In this section we report tests in support of the continuity of density and potential outcomes assumptions, showing there are no discontinuities in the density function of the running variable, in a set of pre-treatment covariates, and at placebo cutoffs (see Figure A.4, Figure A.5, and Figure A.6, respectively).

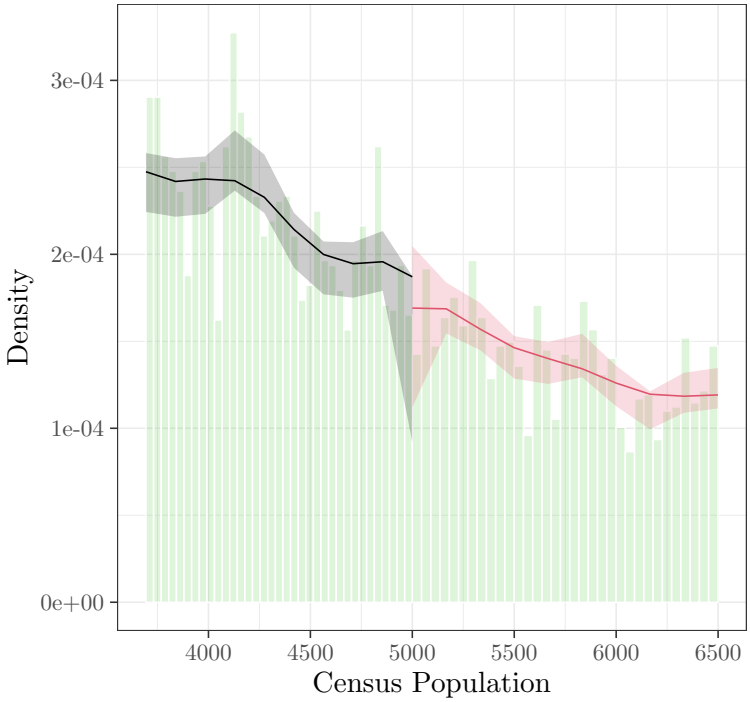

Figure A.4: Manipulation test using the local polynomial density estimators proposed in Cattaneo, Jansson, and Ma (2020). Histogram estimate of the running variable computed with default values in R; local polynomial density estimate (solid dark and red) and robust bias corrected confidence intervals (shaded dark and red) computed using *rddensity* package in R. The number of observations just above the cutoff is not significantly different from the number of observations just below the cutoff (p.value = 0.50).

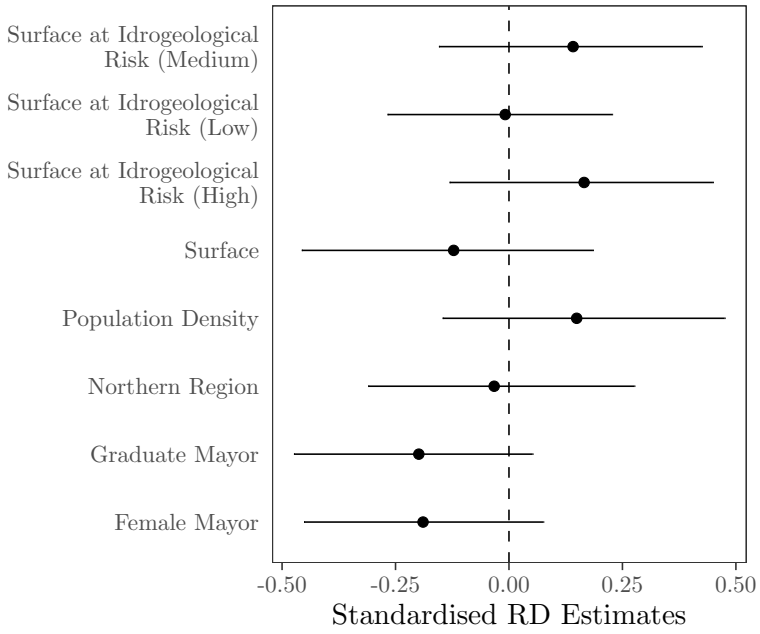

Figure A.5: Standardised RD estimates of the effect on pre-treatment covariates with 95% robust confidence interval Estimates constructed using local polynomial estimators with triangular kernel and CER-optimal bandwidth (as suggested by Cattaneo, Idrobo, and Titiunik 2019, Ch. 5). Robust CI using bias-correction. No covariates included in the estimation.

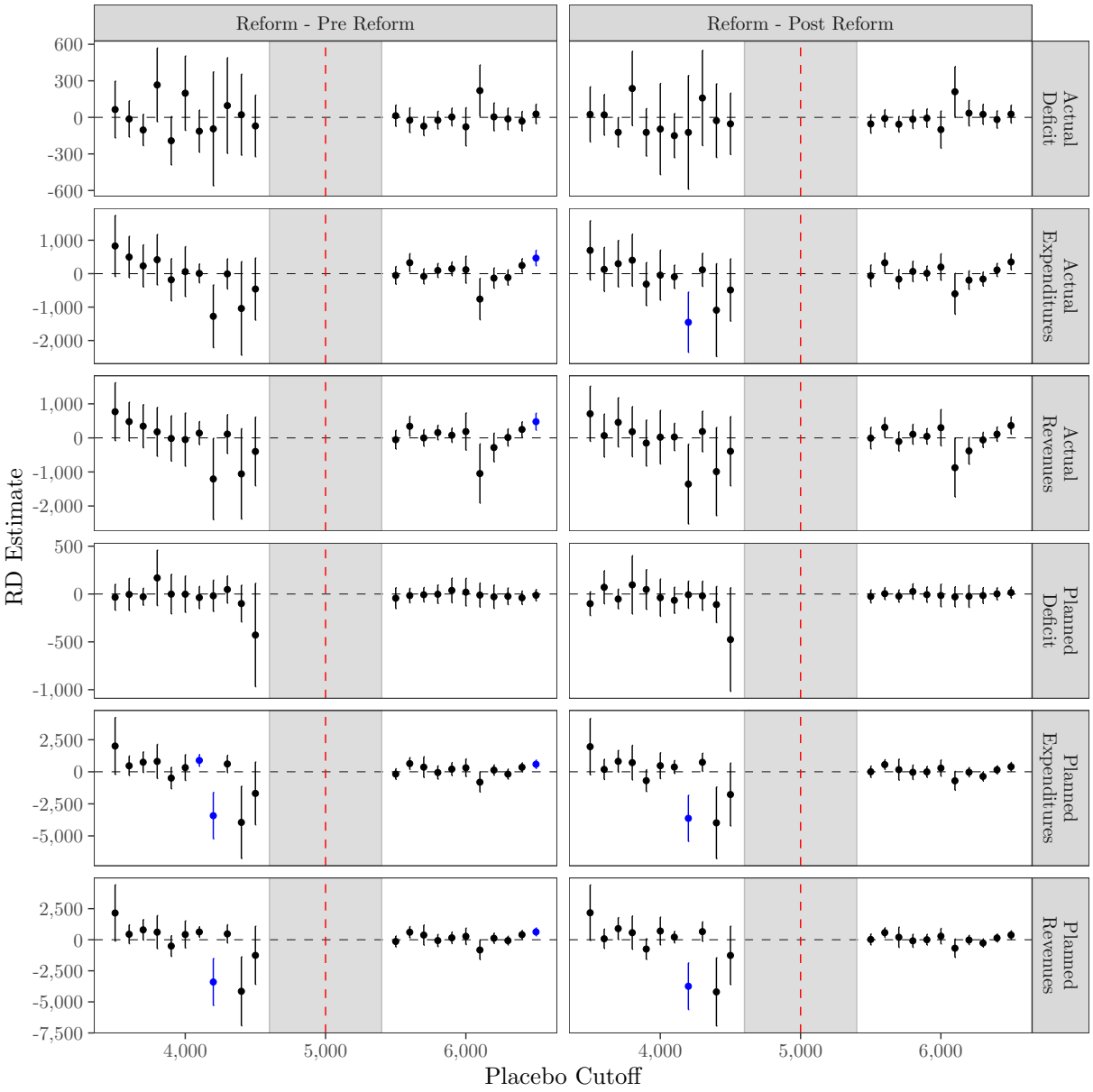

Figure A.6: Diff-in-Disc estimates with 95% CI. Red dashed line at the true cutoff. Blue coefficients when p.value after multiple testing adjustment smaller than 0.05. Multiple-testing adjustment performed separately for each outcome variable with Bonferroni procedure to control for the false discovery rate. Estimates constructed separately on control unit when placebo cutoff < 0, and on treated unit when placebo cutoff > 0. Placebo cutoffs very close to 0 omitted due to small sample size. Estimation performed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Confidence interval constructed using robust standard errors. No covariates included. We fail to detect a discontinuity statistically significant effects in 96.2% of the tests.

A.8 Validity of Diff-in-Disc Estimator

To test the assumption that politicians paid differently do not react differently to a change in the number of politicians, I compare discontinuities at the cutoff before and after a reform that changed the size of government bodies but not on a population threshold basis. Municipalities above and below the 5,000 cutoff – who are paid differently – experienced a decrease in the number of politicians by 20%. If a change in the number of politicians affected the effect of wage treatments on the outcomes, I should detect a significant difference in the discontinuity at the cutoff before and during the placebo reform. Figure A.7 below shows that the difference between the RD estimates in the pre and placebo reform periods is not distinguishable from 0 for all the outcomes. A decrease in the number of politicians therefore does not change the effect of wage policies.

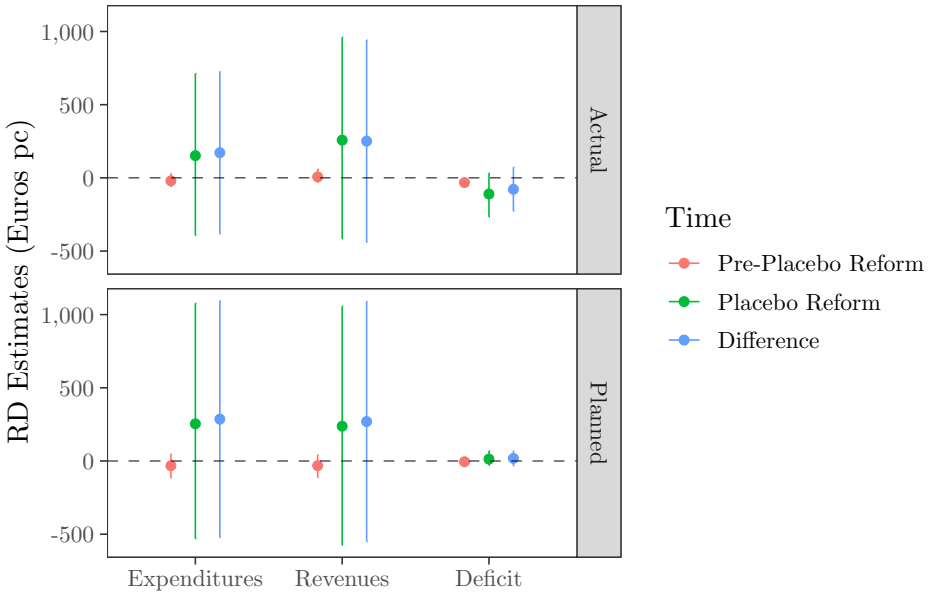

Figure A.7: Diff-in-Disc estimates with 95% CI estimated with alternative bandwidths. Red coefficients estimated with MSE-optimal bandwidth. Underlying RD estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence interval constructed using bias-correction with robust standard errors. Same covariates included in Table A.4.

The "local" parallel trend assumption is indirectly tested in Figure A.8, where we estimate the discontinuities in all the outcomes for every year and show that they are highly stable in the pre reform period. Furthermore, as we show in the *Results* section, the RD effects jump in the reform period and are very similar in the pre- and post-reform periods, suggesting that, absent the reform, changes at the discontinuities do not vary significantly.

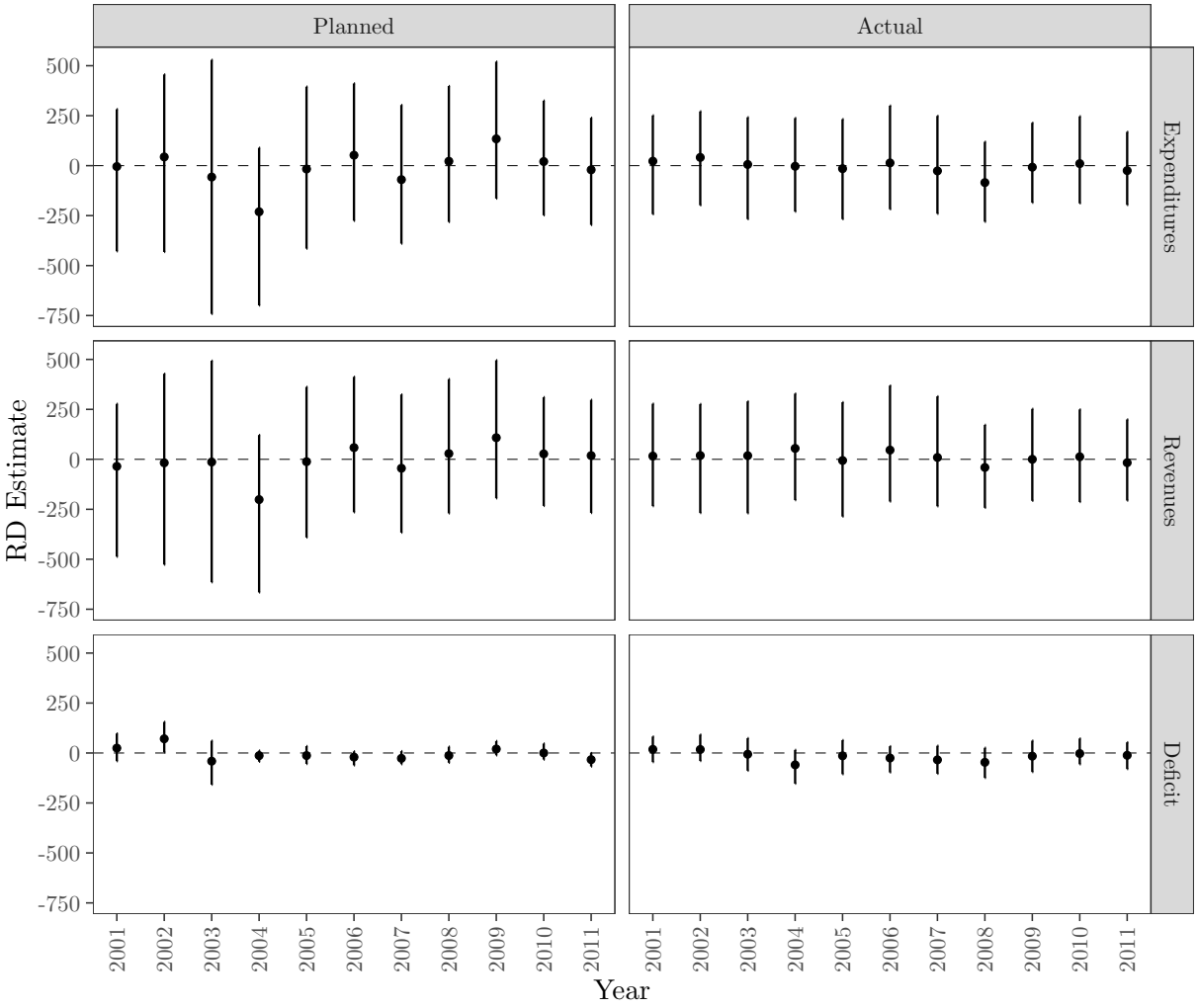

Figure A.8: RD estimates for every year and every outcome in pre-treatment period (before Reform enters into force) Robust 95% confidence intervals, MSE-optimal bandwidth, and triangular kernel. No covariates included.