Polity Size and the Congested Budget: Evidence from Italian Municipalities Online Appendix

Luca Bellodi* Massimo Morelli†

*Hoover Institution and Department of Political Science, Stanford University. bellodi@stanford.edu
†Bocconi University, IGIER, PERICLES, and CEPR. massimo.morelli@unibocconi.it

Online Appendix

A Simple Model for the General Predictions

In this section we present a formal derivation of our main prediction that a larger polity size with fixed bureaucratic capacity reduces the probability of execution of each planned policy.

As in classical models of political agency (see e.g., Fearon, 1999), politicians can always be described as *agents*, in a principal-agent relationship with voters (typically the main principal, although politicians might be agents of interest groups and other external principals). Regardless of which principal an agent is mostly influenced by, the principal–agent relationship always involves a mandate, with respect to which the agent will then be evaluated e.g. for reelection or for any other career incentive. An intuitive characterization of the mandate is a set of policies or programs the politicians is delegated authority to implement. Thus, any politician elected to any elected official role must have a minimum number of projects to propose and try to implement.

Politicians are moved by different goals. The primary concern of an elected official is to look good with voters, who may have to reelect her. Similarly, politicians care about their reputation among the general public, or they might have to satisfy, directly or indirectly, the interest groups that either financed their campaign or are actively lobbying them. Moreover, politicians might have career concerns within the party, or within the electoral coalition, in order to advance their political career to some other political or public administration job once the elected office is over. Activism incentives have been described in the literature for legislators (and hence council members) as much as for members of executive committees, and politicians have been portrayed as agents who seek to please constituents (Mayhew, 1974), interest groups (Rocca and Gordon, 2010), other legislators (Kessler and Krehbiel, 1996), or trying to enhance reputation as reformers or to posture on divisive issues (Gratton et al., 2021; Ash et al., 2017).

Let us now consider two municipalities. In Municipality 1, a polity of k politicians have to agree at time 0 about a set of policies to put on the planned budget, with related financial coverage. On the budget, the set of policies and their respective financial coverage take the form of expenditures and revenues. The set of potential policies is $X \equiv \{x_1, ..., x_n\}$, with $n > k$. For simplicity, let us assume that each policy $x_i, i = 1, n$ has the same tax revenue cost c, and that voters' income and willingness to pay (or government transfers to the municipality) allow a maximum of k projects to be financed, so that k is both the number of elected officials and the maximum number of projects that could be justified or covered by financial resources at the budget planning stage. Municipality 2 differs from Municipality 1 only for polity size, with $h > k$ members. Assume $k > \frac{h-1}{2}$, so that the difference is not too large, and assume that both k and h are odd numbers, in order to avoid tie breaking rules or integer problems. In both municipalities, the decision over which policy to be put on the planned buget is by majority rule, using demand bargaining as clarified below.[1]

Let us assume that each politician has a benefit $B > 0$ in case their most important project gets completed or clearly on the way by the end of the time in office, leaving out the details of whether this benefit B comes from one of the principals mentioned above or from intrinsic motivations of the agent/politician. For simplicity, all other projects would yield zero

[1] Demand bargaining is a simple alternative to the alternating offer protocol in Baron and Ferejohn (1989). See e.g. Morelli (1999).

utility for her. Formally, for each elected official j there exists $x \in X : \ u_j(x) = B$ and $u_j(y) = 0 \ \forall y \in X, y \neq x$.

In line with Morelli (1999), the demand bargaining procedure to reach an agreement on the planned budget works as follows: the members of the polity make a demand, sequentially, and each demand is simply a subset of X. A majority coalition and hence an agreement is formed as soon as a majority of the polity has made compatible demands, i.e., the sum of costs of the demands made by a majority coalition cannot exceed ck. Let us denote by $Y \subseteq X$ the set of policies placed on the planned budget by a committee in equilibrium.

Finally, assume that both municipalities have a bureaucratic capacity constraint and that it is the same for each municipality. Such a bureaucratic capacity constraint can be modeled by assuming that the probability of implementation of any project x in the planned budget is equal to $P(|Y|)$, where $|Y|$ is the number of elements of the set Y and $P(\cdot) \in (0, 1)$ for every positive number of projects, decreasing in the number of projects on the budget. Intuitively, the probability that the bureaucracy manages to implement all the policies is a function of the number of policies on the planned budget (i.e., $|Y|$).

Proposition 1. *(I) If P' decreases not too sharply, the planned budget of Municipality 2 is always larger than the planned budget of Municipality 1.*

(II) Moreover, the ratio of actual over planned budget is lower (in expectation) in Municipality 2 than in Municipality 1, under the same conditions.

Proof. The unique equilibrium of the demand bargaining game in Municipality 1 displays $|Y_1| = \frac{k-1}{2}$, while in Municipality 2 it must be $|Y_2| = \frac{h-1}{2} > |Y_1|$. In both, a simple majority of members each obtains the planning of her own preferred project. Whatever the order in which elected officials move in the demand game, where the order could come from any institutional or randomization protocol, the first $\frac{l-1}{2}$ members – $l = k, h$ – have each incentive to demand their own preferred project, since the probability of completion is positive and any subset of X is justifiable by a revenue coverage by assumption. Even though the addition of one demand reduces the probability of completion P of all projects marginally, the deviation to not demand such a project would yield zero. The unique equilibrium is robust to changes in the utility function, for example to assuming that each member also would derive positive utility from the projects of other members in the majority coalition, as long as at least for some members such a utility from projects of others is not too large.

Part (II) follows mechanically.

QED.

B Data Sources

To assemble the dataset, we rely on four sources of data.

1. We obtained full lists of municipalities and unique budget identifiers directly from the Local Public Finance Directorate of the Ministry of Interior. We then used the unique IDs to build URLs and scrape budget data from the on-line ministerial repository. This repository, available at `https://finanzalocale.interno.gov.it/apps/floc.php/in/cod/4` reports the data contained in the certificates of the balance sheets that the municipalities, provinces, and metropolitan cities *must* transmit to the Ministry of Economy and Finance, pursuant to Legislative Decree 118/2011. Every municipality has to communicate this data and hence non-compliance and missing data are not a concern for the definition of the sample of municipalities.

2. We merged budget data with data on local government composition (composition and biographical information on mayors and municipal politicians) obtained from the Database on Local and Regional Administrators, curated by the Ministry of the Interior. The data can be accessed at `https://dait.interno.gov.it/elezioni/open-data`. We exact matched the two datasets based on the name of the municipalities. We manually checked those unmatched municipalities and we resolved conflicts on a case-by-case basis. The Database on Local and Regional Administrators contains another unique ID for municipalities which is the one produced by the National Institute of Statistics, hence we merge all other data sources by this ID.

3. We obtained data on socio-demographic and territorial characteristics of municipalities as well as the indicators on spending and collection capacity, and the share of bureaucrats with a university degree from the National Institute of Statistics. Data available at `http://dati.statistiche-pa.it/`.

4. We accessed data on the number of municipal employees from the national account of the Italian General Accounting Office. Data available at `contoannuale.mef.gov.it`. This data does not report the unique IDs of the National Institute of Statistics, hence the dataset has been exact matched on the name of the municipalities.

5. We obtained data on the personal income certificates of municipal residents from the Ministry of the Economy and Finance (the Italian acronym of this certificate is IRPEF). Data available at `https://www.finanze.gov.it/it/statistiche-fiscali/`. We merged this data by the unique ID of the National Institute of Statistics.

C Sample of Municipalities and Inflation Adjustment of Budget Data

From the total sample of municipalities, we removed the municipalities located in the five special statute regions (Sicilia, Sardegna, Valle d'Aosta, Friuli-Venezia Giulia, and Trentino-Alto Adige/Südtirol) for the threshold mechanisms 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 remaining 15 ordinary-statute regions.

All the outcome variables are per capita and adjusted to the 2018 inflation level. Figures are therefore expressed in real terms and are comparable over time. We applied the consumer price index for currency evaluations as of January of every year (Jan. 2018 index = 100). Time-series downloaded from the archive of the National Institute of Statistics, `www.istat.it/it/archivio/30440`.

D Policies Based on Population Thresholds

The tables below report rules which change discontinuously at population thresholds. We also report the legal sources, the time frame when the rules remained in force, and the various cutoffs at which they operated. In particular, Table D.2 shows the wage policies changing at the 5,000 cutoff.

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 Members Executive Committee	2001-Present	Art. 4, Min. Dec. 119/2000	1, 5, 50
Reimbursement to Local Councillors	1998-2000 2001-Present	Art. 10, Law 861/1985 Art. 1 and Table A, Min. Dec. 119/2000	30 1, 10, 30
Wage President of Council	2001-Present	Art. 5, Min. Dec. 119/2000	1, 15
Presence of Neighborhood Councils	1998-2000 2001-Present	Art. 13, Law 142/90 Art. 17, Leg. Dec. 267/2000	30 (optional)
Number of Financial 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
Presence of 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 on Party Lists	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

Table D.1: Legal sources of all rules/policies 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 Center						?	
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 D.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 center and neighborhood councils are local councils that can be established in any neighborhood. Two important exceptions: wage/reimbursement thresholds at 1,000 were introduced only in 2001.

E Reforms to the Size of Municipal Government Bodies

In 2011, two reforms were passed aimed at reducing the number of politicians in municipal governments, with the goal of controlling public expenditures. The first reform (Reform 1) affected all municipalities, without affecting the population-threshold mechanism. A second reform (Reform 2), the one studied in this paper, introduced a new threshold of 5,000 inhabitants, which was then repealed in 2014. Table E.3 shows how the size of government bodies changed under these two different reforms.

			Election Years 2011-2013					
Threshold	**Pre-Reform**		**Reform 1**		**Reform 2**		**Post-Reform**	
[1,000]	Council	Ex. C.	Council	Ex. C.	Council	Ex. C.	Council	Ex. C.
0 – 3	12	4	10	3	6	2	10	2
3 – 5	16	6	13	4	7	3	12	4
5 – 10					10	4		
10 – 30	20	7	16	5	16	5	16	5
> 30	30	10	24	7	24	7	24	7

Table E.3: Number of local councilors and cap to size of executive committees before and after two reforms which were passed in 2011.

F Descriptive Statistics

Table F.4 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	15.2	16.0	1.9
Members of Exec. Comm.	3.2	3.0	1.7	3.9	4.0	1.1
Planned Budget						
Expenditures pc	1,769.0	1,320.9	2,182.7	1,309.3	1,118.3	831.5
Revenues pc	1,754.1	1,306.4	1,974.6	1,303.6	1,108.4	849.9
Deficit pc	12.3	5.7	251.3	5.7	5.3	200.5
Actual Budget						
Expenditures pc	999.8	809.3	977.9	804.3	704.2	478.3
Revenues pc	1,035.8	825.1	1,104.6	826.4	723.2	485.7
Deficit pc	-36.1	-13.3	454.6	-22.1	-16.4	190.0

Table F.4: 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.

G RD Plots

In the figures below we report RD plots with WLS fitted lines estimated separately within the MSE-optimal bandwidth (vertical dotted line) above and below the 5,000 cutoff for planned (Figure G.1) and actual budget figures (Figure G.2).

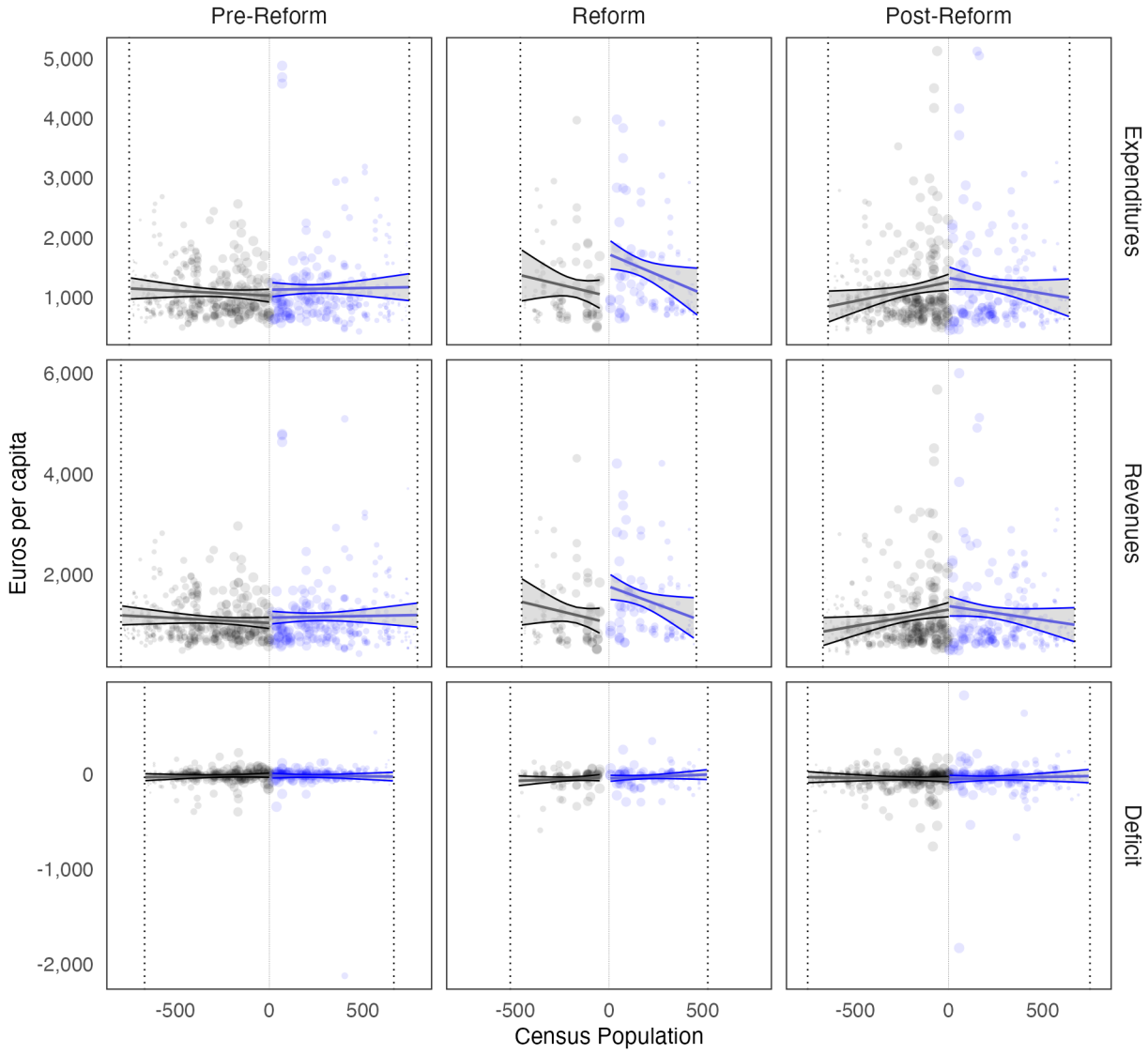

Figure G.1: RD plot with fitted WLS line estimated separately above and below the cutoff for planned budget. Each dot is a municipality-year observation, and the size of the dot is a function of the weight determined by the triangular kernel function based on the ratio of the distance of each observation from the cutoff and the MSE-minimizing bandwidth (vertical dotted line).

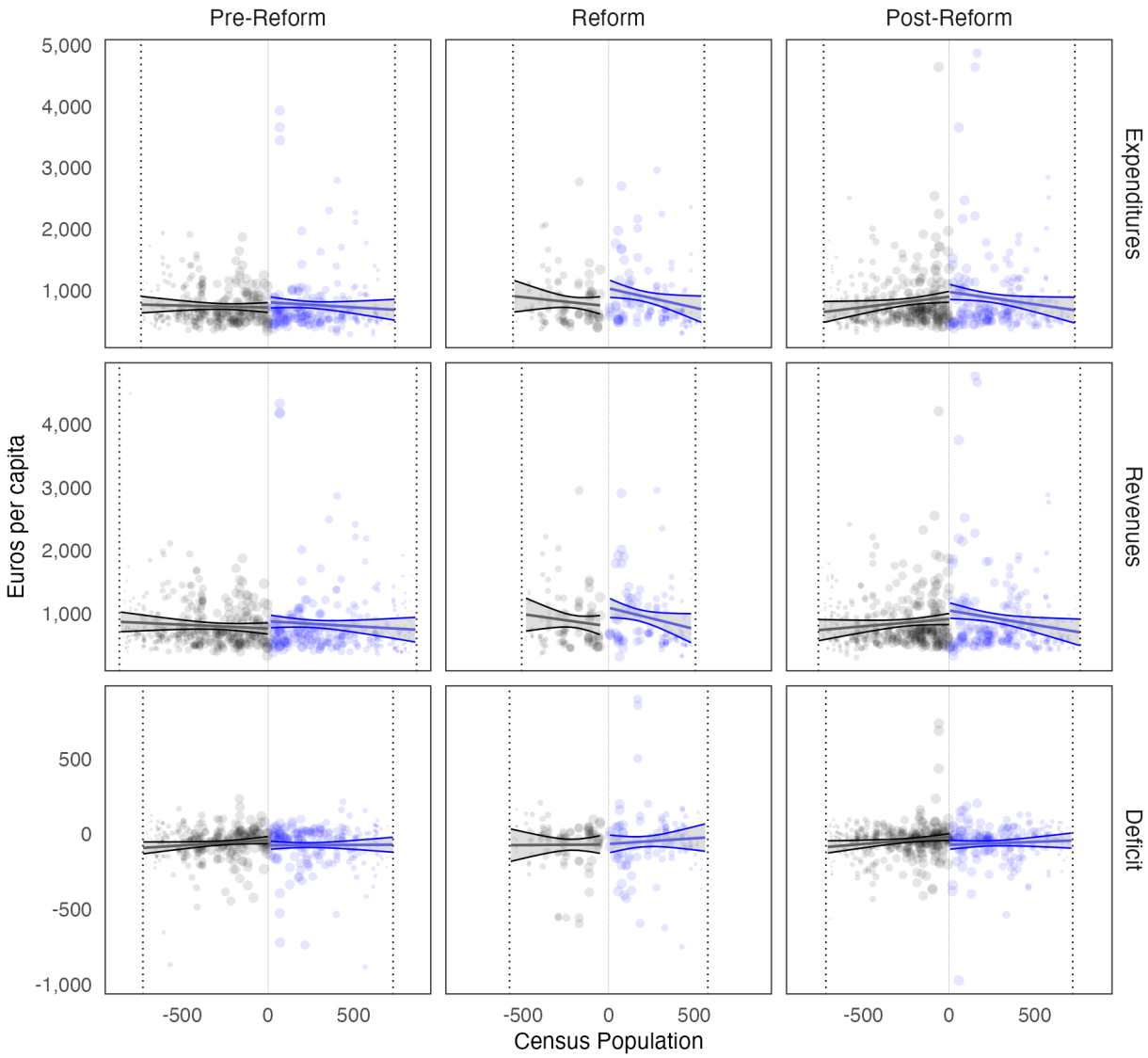

Figure G.2: RD plot with fitted WLS line estimated separately above and below the cutoff for actual budget. Each dot is a municipality-year observation, and the size of the dot is a function of the weight determined by the triangular kernel function based on the ratio of the distance of each observation from the cutoff and the MSE-minimizing bandwidth (vertical dotted line).

H Regression Tables

H1 Main Analysis

In the table below we report the RD results for the three time periods separately. These are the estimates reported in the top panels of Figure 2 in the main text. The analysis was performed with the *rdrobust* package in R (Calonico et al., 2015).

	Planned Budget			**Actual Budget**		
Outcome	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Pre-Reform						
Estimate	72.8	142.2	-51.2	147.3	159.1	6.1
SE	(96.1)	(136.9)	(30.8)	(145.6)	(141.9)	(25.7)
p.value	0.339	0.197	0.075	0.239	0.204	0.710
h	1043.8	708.1	627.3	780.6	817.4	500.3
Obs. Used	751	456	400	515	549	332
Reform						
Estimate	197.4	238.5	-5.7	764.7	815.9	5.3
SE	(125.1)	(146.8)	(43.7)	(222.9)	(235)	(33.6)
p.value	0.080	0.053	0.811	0.000	0.000	0.763
h	845.7	660.7	598.2	556.4	492.6	453.1
Obs. Used	339	255	222	201	181	167
Post-Reform						
Estimate	139.7	179.6	-41.7	153.7	133.3	11.8
SE	(143.3)	(148.8)	(34.6)	(194)	(218.4)	(56)
p.value	0.408	0.265	0.219	0.497	0.632	0.721
h	631.7	656.2	766.5	573.0	591.5	734.4
Obs. Used	456	483	576	425	435	551

Table H.5: 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 (h). Robust p.values computed using bias-correction with robust standard errors. Covariates include: population density, surface (sq.km), surface at low, medium, and high hydro-geological risk (sq.km) – all log transformed –, gender, mayor with university degree (dummy), white-collar mayor (dummy), year and province dummies.

In the table below we report the diff-in-disc estimates comparing the RD estimates in $T = 1$ (under the reform) with the RD estimates in $T \in \{0, 2\}$, pre- and post-reform, respectively. These are the estimates reported in the bottom panels of Figure 2 in the main text. The point estimate is the difference between the RD point estimates in the two periods, and the standard error of the difference has been computed with the following formula, as in Klašnja and Titiunik (2017): $SE_{DD} = \sqrt{SE^2_{RD|T=1} + SE^2_{RD|T\in\{0,2\}}}$, where $T = 1$ refers to the SE of the RD point estimate for reform period, and $T \in \{0, 2\}$ for pre- and post-reform periods, respectively.

	Reform - Pre-Rerom			**Reform - Post-Reform**		
Outcome	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Planned Budget						
Diff.	617.5	656.8	-0.8	611.0	682.6	-6.5
SE	(266.2)	(274.5)	(42.3)	(295.5)	(320.8)	(65.3)
p.value	0.020	0.017	0.986	0.039	0.033	0.921
Actual Budget						
Diff.	57.6	58.9	36.0	57.6	58.9	36.0
SE	(190.2)	(209.0)	(55.8)	(190.2)	(209.0)	(55.8)
p.value	0.762	0.778	0.519	0.762	0.778	0.519

Table H.6: Diff-in-Disc estimates (i.e., difference in RD point estimates reported in Table A.5) as displayed in Figure 2 for every outcome and planned and actual budgets.

H2 Mechanism

In the tables below we report the RD estimates in the two samples of municipalities whose capacity indicator is below and above the median value. We use two outcome variables produced by the National Institute of Statistics.

- Collection capacity is an indicator computed as the ratio of actual over planned revenues.
- Spending capacity is an indicator computed as the ratio of actual over planned expenditures.

For each outcome, we first report RD estimates in the two samples of municipalities above and below the median capacity value (Table H.7 and Table H.9) and then the diff-in-disc estimates in the two samples as well as the difference in diff-in-disc estimates between the samples (Table H.8 and Table H.10).

Moreover, we replicate this analysis using collection and spending capacity indicators built from the budget data we scraped ourselves and show the results are similar.

Capacity Indicators Produced by the National Institute of Statistics

	DV: Collection Capacity (Actual/Planned Revenues)					
	Below Median			Above Median		
Time Period	Pre-Reform	Reform	Post-Reform	Pre-Reform	Reform	Post-Reform
Capacity Indicator: % Bureaucrats with Degree						
Estimate	0.00	-0.29	0.04	-0.06	-0.04	0.02
SE	(0.05)	(0.06)	(0.05)	(0.03)	(0.06)	(0.03)
p.value	0.98	0.00	0.30	0.03	0.46	0.78
h	525.13	434.07	561.69	452.23	744.38	576.80
Obs. Used	167	84	199	158	128	210
Capacity Indicator: N. Bureaucrats						
Estimate	-0.06	-0.11	0.04	-0.02	-0.09	0.03
SE	(0.03)	(0.07)	(0.03)	(0.04)	(0.09)	(0.05)
p.value	0.01	0.04	0.17	0.78	0.33	0.49
h	453.71	531.25	460.28	592.63	509.78	705.23
Obs. Used	175	115	247	158	66	139

Table H.7: RD estimates for each time period and each outcome in two samples consisting of municipalities whose capacity indicator is above and below the median. Outcome is collection capacity indicator produced by National Institute of Statistics, i.e., the ratio of actual over planned revenues. Variable proxying capacity indicator has been reported in each panels. Estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth (h). Robust p.values computed using bias-correction with robust standard errors. Same covariates used in main analysis.

	DV: Collection Capacity (Actual/Planned Revenues)					
	Below Median		Above Median		Above - Below	
Time Period	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform
Capacity Indicator: % Bureaucrats with Degree						
Difference	-0.290	-0.330	0.020	-0.060	0.310	0.270
SE	(0.08)	(0.08)	(0.07)	(0.07)	(0.11)	(0.11)
p.value	0.00	0.00	0.82	0.44	0.00	0.01
Capacity Indicator: N. Bureaucrats						
Difference	-0.060	-0.150	-0.060	-0.110	-0.010	0.030
SE	(0.08)	(0.08)	(0.10)	(0.10)	(0.12)	(0.13)
p.value	0.46	0.05	0.51	0.27	0.95	0.79

Table H.8: Diff-in-Disc estimates computed separately for below- and above-median samples. Outcome variable is collection capacity indicator produced by National Institute of Statistics, i.e., the ratio of actual and planned revenues. Same covariates used in main analysis. 'Above - Below' columns report the difference in the diff-in-disc estimates, with the standard error calculated with the following formula: $SE = \sqrt{SE^2_{Above} + SE^2_{Below}}$, where *Above* refers to the SE of the diff-in-disc point estimate for the above-median sample, and *Below* refers to the SE of the diff-in-disc point estimate for the below-median sample.

	DV: Spending Capacity (Actual/Planned Expenditures)					
	Below Median			Above Median		
Time Period	Pre-Reform	Reform	Post-Reform	Pre-Reform	Reform	Post-Reform
Capacity Indicator: % Bureaucrats with Degree						
Estimate	-0.03	-0.26	-0.01	-0.03	-0.06	-0.02
SE	(0.04)	(0.06)	(0.04)	(0.03)	(0.08)	(0.03)
p.value	0.58	0.00	0.82	0.13	0.56	0.61
h	606.43	426.70	829.33	561.52	643.10	600.65
Obs. Used	179	83	301	179	109	215
Capacity Indicator: N. Bureaucrats						
Estimate	-0.04	-0.10	0.02	-0.06	-0.12	-0.07
SE	(0.03)	(0.08)	(0.03)	(0.04)	(0.08)	(0.06)
p.value	0.11	0.09	0.40	0.18	0.12	0.15
h	824.67	489.59	520.61	597.04	436.88	663.41
Obs. Used	323	112	277	158	59	126

Table H.9: RD estimates for each time period and each outcome in two samples consisting of municipalities whose capacity indicator is above and below the median. Outcome is spending capacity indicator produced by National Institute of Statistics, i.e., the ratio of actual and planned expenditures per capita. Variable proxying capacity indicator has been reported in each panels. Variable proxying capacity indicator has been reported in each panels. Estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth (h). Robust p.values computed using bias-correction with robust standard errors. Same covariates used in main analysis.

	DV: Spending Capacity (Actual/Planned Expenditures)					
	Below Median		Above Median		Above - Below	
Time Period	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform
Capacity Indicator: % Bureaucrats with Degree						
Difference	-0.230	-0.240	-0.030	-0.050	0.200	0.200
SE	(0.07)	(0.07)	(0.08)	(0.08)	(0.11)	(0.11)
p.value	0.00	0.00	0.73	0.58	0.07	0.08
Capacity Indicator: N. Bureaucrats						
Difference	-0.050	-0.120	-0.060	-0.040	0.000	0.080
SE	(0.08)	(0.08)	(0.08)	(0.09)	(0.12)	(0.13)
p.value	0.51	0.16	0.48	0.65	0.97	0.55

Table H.10: Diff-in-Disc estimates computed separately for below- and above-median samples. Outcome variable is spending capacity indicator produced by National Institute of Statistics, i.e., the ratio of actual and planned expenditures per capita. Same covariates used in main analysis. 'Above - Below' columns report the difference in the diff-in-disc estimates, with the standard error calculated with the following formula: $SE = \sqrt{SE^2_{Above} + SE^2_{Below}}$, where *Above* refers to the SE of the diff-in-disc point estimate for the above-median sample, and *Below* refers to the SE of the diff-in-disc point estimate for the below-median sample.

Capacity Indicators Produced from Scraped Budget Data

In Tables H.11 and H.12 we replicate the same analysis this time building the measures of collection and spending capacity from the scraped budget data. Consistent with the measures produced by the National Institute of Statistics, we compute the two indicators as the ratio of actual expenditures (revenues) per capita over planned expenditures (revenues) per capita. These indicators are identical to those produced by the National Institute of Statistics, except for some missingness in both data sources. For instance, for some municipality-year pairs there is no available budget data but the National Institute of Statistics was still able to produce spending and collection capacity indicators, and vice versa.

The results are similar except for the confidence intervals of the effect of spending capacity, which are larger (p.values = .24 for the reform – pre-reform period and .28 for the reform–post-reform period).

	DV: Collection Capacity (Actual/Planned Revenues)					
	Below Median		**Above Median**		**Above - Below**	
Time Period	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform
Capacity Indicator: % Bureaucrats with Degree						
Difference	-0.27	-0.28	0.02	-0.05	0.29	0.23
SE	(0.08)	(0.08)	(0.07)	(0.07)	(0.11)	(0.11)
p.value	0.00	0.00	0.78	0.45	0.01	0.03
Capacity Indicator: N. Bureaucrats						
Difference	-0.05	-0.14	-0.06	-0.11	-0.01	0.03
SE	(0.08)	(0.08)	(0.11)	(0.12)	(0.14)	(0.14)
p.value	0.48	0.06	0.59	0.34	0.96	0.84

Table H.11: Diff-in-Disc estimates computed separately for below- and above-median samples. Outcome variable is collection capacity computed from scraped budget data as the ratio of actual and planned revenues per capita. Same covariates used in main analysis. 'Above - Below' columns report the difference in the diff-in-disc estimates, with the standard error calculated with the following formula: $SE = \sqrt{SE^2_{Above} + SE^2_{Below}}$, where *Above* refers to the SE of the diff-in-disc point estimate for the above-median sample, and *Below* refers to the SE of the diff-in-disc point estimate for the below-median sample.

	DV: Spending Capacity (Actual/Planned Expenditures)					
	Below Median		**Above Median**		**Above - Below**	
Time Period	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform	Reform - Pre-Reform	Reform - Post-Reform
Capacity Indicator: % Bureaucrats with Degree						
Difference	-0.16	-0.17	-0.02	-0.05	0.14	0.13
SE	(0.08)	(0.08)	(0.08)	(0.09)	(0.12)	(0.12)
p.value	0.04	0.02	0.78	0.59	0.24	0.28
Capacity Indicator: N. Bureaucrats						
Difference	-0.05	-0.11	-0.03	-0.01	0.03	0.11
SE	(0.08)	(0.08)	(0.09)	(0.10)	(0.12)	(0.13)
p.value	0.52	0.17	0.76	0.92	0.83	0.42

Table H.12: Diff-in-Disc estimates computed separately for below- and above-median samples. Outcome variable is spending capacity computed from scraped budget data as the ratio of actual and planned expenditures per capita. Same covariates used in main analysis. 'Above - Below' columns report the difference in the diff-in-disc estimates, with the standard error calculated with the following formula: $SE = \sqrt{SE^2_{Above} + SE^2_{Below}}$, where *Above* refers to the SE of the diff-in-disc point estimate for the above-median sample, and *Below* refers to the SE of the diff-in-disc point estimate for the below-median sample.

In the figure below we report the RD and Diff-in-Disc results reported in Figure 2 on different sub-samples of municipalities whose measures of bureaucratic capacity are above and below the median. It is clear from the top-left panel that the reform had an effect on actual budget for municipalities with a share of bureaucrats with university degree above the median value.

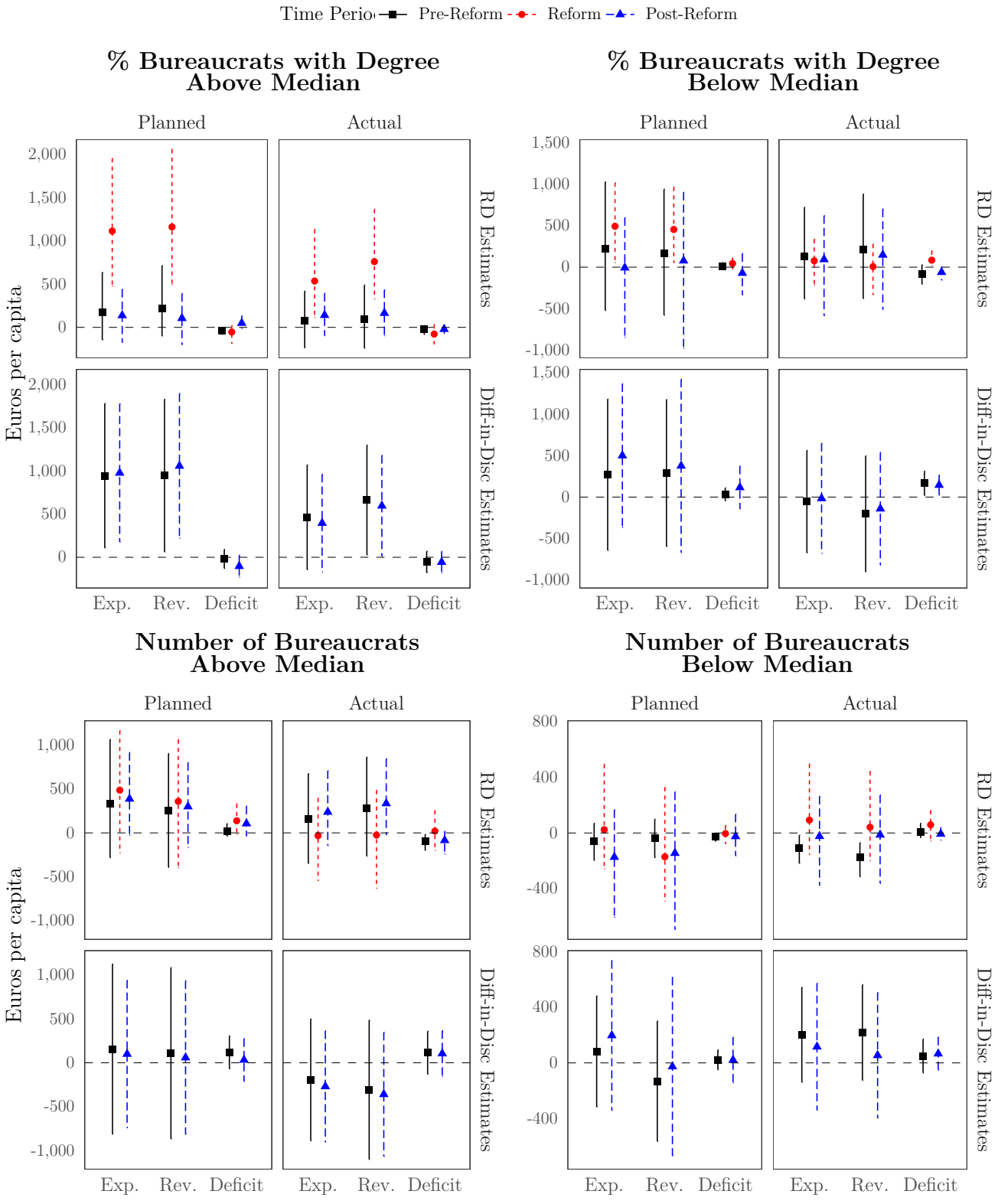

Figure H.3: RD and Diff-in-Disc estimates with 95% robust confidence intervals estimated on different samples of municipalities above and below the two measures of bureaucratic capacity. Same covariates included in the main analysis.

I Robustness Tests

I1 Removing Covariates

In the tables below we show the results are robust to omitting the covariates from the estimation (Table I.13).

	Reform - Pre-Rerom			**Reform - Post-Reform**		
	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Planned Budget						
Difference	606.7	618.9	-7.2	635.3	651.3	-13.2
SE	(305.0)	(317.6)	(38.7)	(327.7)	(356.1)	(64.6)
p.value	0.047	0.051	0.853	0.053	0.067	0.838
Actual Budget						
Difference	209.2	160.0	55.2	209.2	160.0	55.2
SE	(209.7)	(216.3)	(59.7)	(209.7)	(216.3)	(59.7)
p.value	0.318	0.459	0.355	0.318	0.459	0.355

Table I.13: Diff-in-Disc estimates for every outcome and planned and actual budgets without including covariates.

I2 Additional Covariates

In the tables below we show the results are robust to including a larger set of covariates (Table I.14).

	Reform - Pre-Rerom			**Reform - Post-Reform**		
	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Planned Budget						
Difference	496.6	509.0	2.1	519.4	539.4	-4.3
SE	(267.5)	(268.4)	(41.6)	(284.1)	(306.6)	(63.8)
p.value	0.063	0.058	0.959	0.068	0.079	0.947
Actual Budget						
Difference	64.8	40.4	50.9	64.8	40.4	50.9
SE	(194.0)	(205.0)	(55.0)	(194.0)	(205.0)	(55.0)
p.value	0.739	0.844	0.355	0.739	0.844	0.355

Table I.14: Diff-in-Disc estimates for every outcome and planned and actual budgets estimated including a larger set of covariates: 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), white-collar mayor (dummy), right-wing mayor (dummy), left-wing mayor (dummy), average personal income declared by municipal residents, province and year fixed effects.

I3 Alternative Bandwidths

Figure I.4 below shows the diff-in-disc estimates are robust to using using alternative bandwidths.

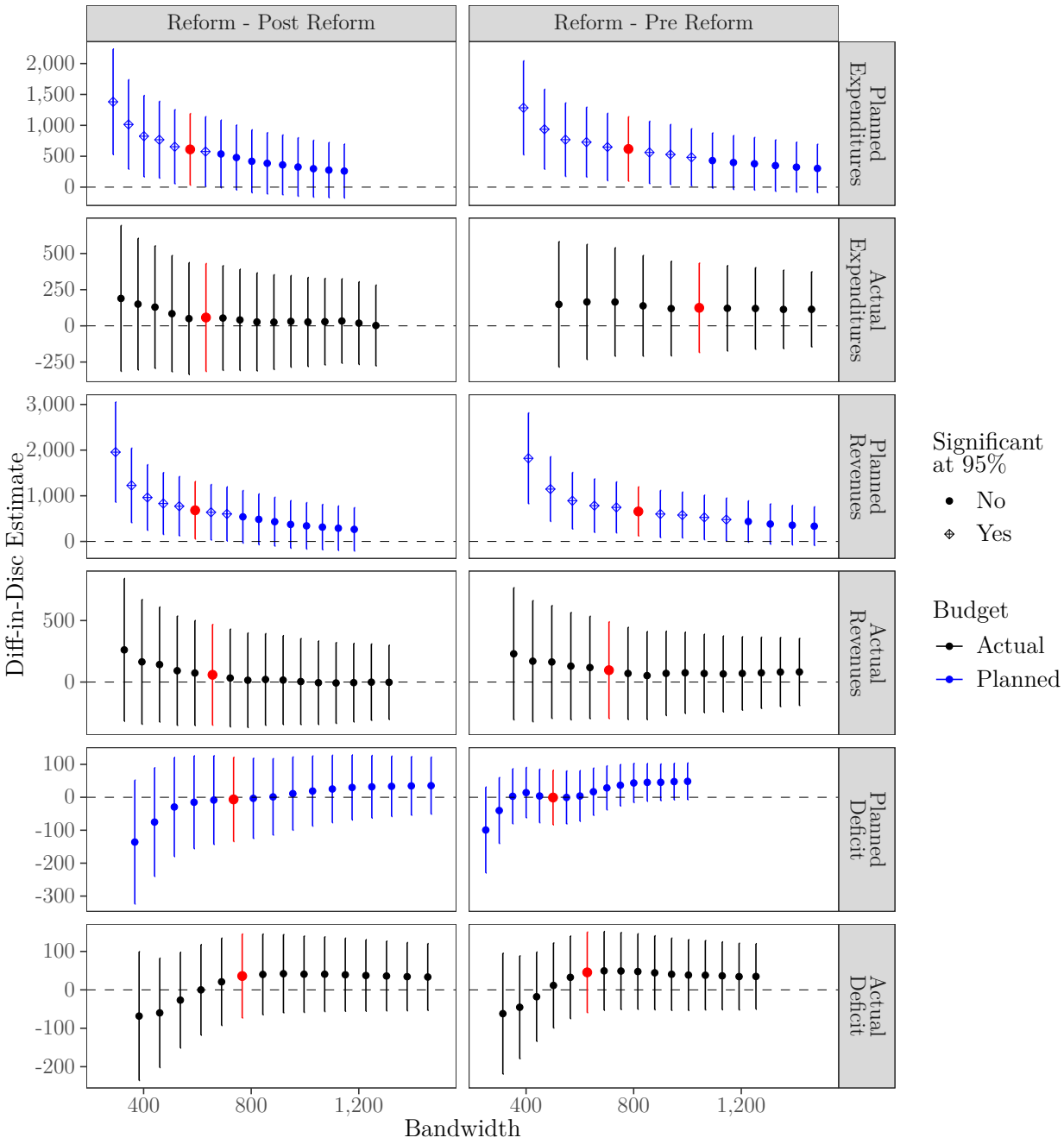

Figure I.4: Diff-in-Disc estimates with 95% robust confidence intervals estimated with alternative bandwidths. Red coefficients estimated with MSE-optimal bandwidth. Underlying RD estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence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

I4 Additional Outcomes from National Institute of Statistics

To show the results are robust to the selection of outcomes, in the table below we replicate the main analysis using two indicators built by the National Institute of Statistics capturing collection and spending capacity and remainder of administration for the same sample of municipalities. The decreased spending and collection capacity and the imprecise estimate for deficit strengthen the main results for which treated municipalities during the reform increase planned spending and revenues (the gap between planned and actual widens, as the effects for spending and collection capacity indicate) but these are not implemented in practice (no statistically significant effect on deficit).

	Reform - Pre-Rerom			**Reform - Post-Reform**		
	Spending Capacity	Collection Capacity	Deficit	Spending Capacity	Collection Capacity	Deficit
Difference	-0.16	-0.15	0.03	-0.15	-0.18	-0.29
SE	(0.06)	(0.06)	(0.14)	(0.07)	(0.07)	(0.30)
p.value	0.015	0.016	0.832	0.022	0.007	0.327

Table I.15: Diff-in-Disc estimates using alternative outcomes built by 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. No covariates included.

I5 Government Term and Fiscal Cycles

Because the analysis is performed on a sample covering three calendar years, pre- and post-reform differences might be confounded by the year-of-term effect, with municipalities more ahead in the government term (and closer to new elections) more likely to spend more. We address this in two ways:

- We show the results are robust to including year-of-term dummies as a covariate (see Table I.16).
- We show that the estimates after removing the effect of the government term cycle on the outcomes are even larger and more precisely estimated. To partial out the effect of the government term on budget data we de-trend the outcomes by taking the residuals of a regression of each outcome on government year-of-term dummies and use these as outcomes in the analysis (see Table I.17). We report the outcome variables expressed in nominal terms and after being detrended in Figure I.5 below.

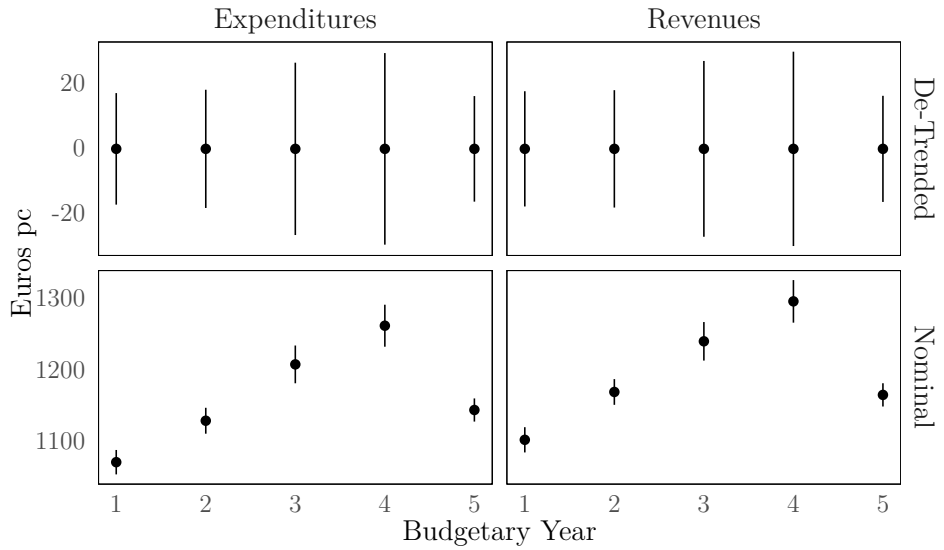

Figure I.5: Average expenditures and revenues per capita over the government term. Top panel reports de-trended outcomes, whereas bottom panel reports nominal outcomes. De-trended outcomes are the residual of a linear regression regressing the nominal outcome on year-of-term dummies (to partial out the effect of business cycles non-parametrically). Panels on the bottom show nominal averages.

	Reform - Pre-Rerom			Reform - Post-Reform		
Outcome	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Planned Budget						
Diff.	649.0	711.5	-0.7	663.5	755.3	-6.5
SE	(268.9)	(276.7)	(42.3)	(299.2)	(324.3)	(65.4)
p.value	0.016	0.010	0.987	0.027	0.020	0.921
Actual Budget						
Diff.	92.0	82.9	28.2	92.0	82.9	28.2
SE	(194.1)	(209.1)	(54.6)	(194.1)	(209.1)	(54.6)
p.value	0.636	0.692	0.606	0.636	0.692	0.606

Table I.16: Diff-in-Disc estimates for each outcome and planned and actual budgets. Same covariates as in main analysis with the addition of year-of-term dummies.

	Reform - Pre-Rerom			Reform - Post-Reform		
	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Planned Budget						
Difference	655.5	697.6	-2.2	649.5	730.6	-6.5
SE	(273.4)	(275.8)	(42.6)	(301.0)	(324.3)	(65.6)
p.value	0.016	0.011	0.959	0.031	0.024	0.922
Actual Budget						
Difference	76.6	63.7	34.6	76.6	63.7	34.6
SE	(194.0)	(209.0)	(55.0)	(194.0)	(209.0)	(55.0)
p.value	0.693	0.761	0.529	0.693	0.761	0.529

Table I.17: Diff-in-Disc estimates. Outcomes from RD estimates are the residuals of linear regressions of each outcome on the year-of-term variable. Same covariates included in main analysis.

I6 Gender Quotas

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 councilors in the government. This has no support in the data, as evidenced by the similar effect size of the diff-in-disc estimates in the pre- and post-reform period, which suggests that the gender composition of the local council has a negligible effect on local public finance. However, in the pre-gender quotas period (earlier than 26 December 2012), many municipalities already renewed their government bodies under the reform and became treated before gender quotas entered into force. Therefore, by limiting the analysis to municipalities that held elections before gender quotas were introduced, we can isolate the effect of having more politicians alone.

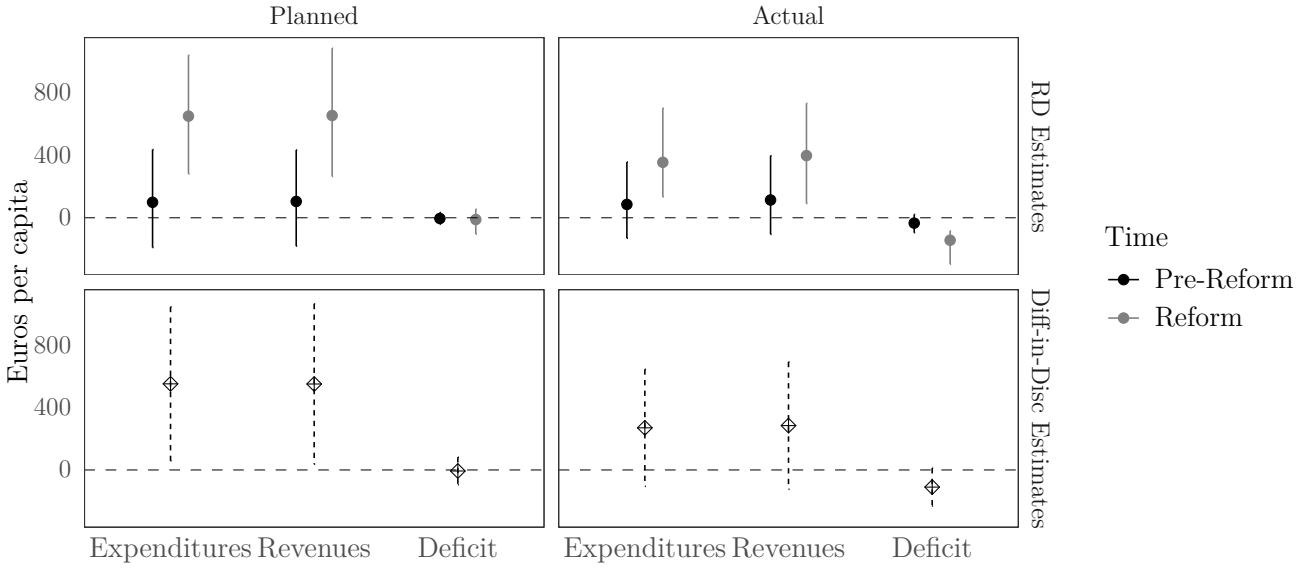

Figure I.6: RD and Diff-in-Disc estimates from sample of municipalities which held elections earlier than 26 December 2012, when the gender quotas entered into force for municipalities above the 5,000 population threshold. Estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence intervals constructed using bias correction with robust standard errors. No covariates included.

I7 Types of Revenues

In Figure I.7 we show the results for planned and actual revenues per capita are not driven by one particular type of revenues.

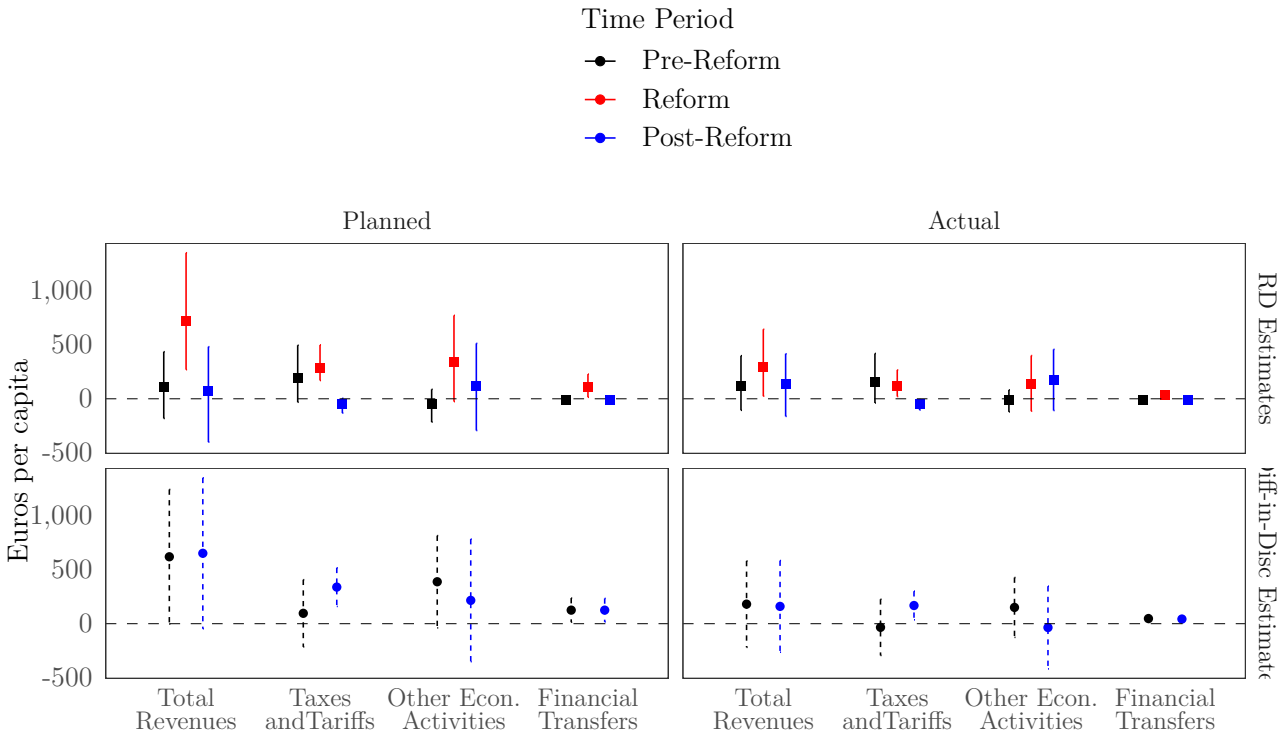

Figure I.7: RD and Diff-in-Disc estimates and robust 95% confidence intervals. Outcomes are revenues per capita and the three components thereof: taxes and tariffs, financial transfers from higher levels of government, and other economic activities of the municipalities. No covariates included.

I8 Selection Effect of the Reform

Table I.18 reports the effect of the reform on the share of councilors and members of the executive committee with a university degree.

	Reform - Pre-Rerom		**Reform - Post-Reform**	
Outcome	% Councillors with Degree	% Members of Ex. Comm. with Degree	% Councillors with Degree	% Members of Ex. Comm. with Degree
Difference	0.12	0.06	0.15	0.03
SE	(0.04)	(0.02)	(0.04)	(0.02)
p.value	0.001	0.000	0.000	0.089

Table I.18: Diff-in-Disc estimates of the effect of the reform on the share of councillors and members of the executive committee with a university degree. Same covariates included in main analysis.

I9 Diff-in-Disc Estimated with Single Equation

In Table I.19 below we show the diff-in-disc estimate by estimating a full equation with two time periods (municipalities holding elections under the reform and not, as in Grembi et al. (2016). We fit local WLS models separately on the observations above and below the cutoff and for municipalities holding elections when the reform was into force ($T = 1$) and not ($T \in \{0, 2\}$). Weights are determined by the triangular kernel function based on the ratio between the distance of unit i from the cutoff and the MSE minimizing bandwidth. We estimate the optimal bandwidth pooling all time periods but results are robust to estimating two different bandwidths in $T = 1$ and $T \in \{0, 2\}$ and then averaging the two. Units outside the optimal bandwidth receive a weight equal to zero. We then estimate the following equation:

$$\begin{aligned} Y_{it} = {} & \delta_0 + \delta_1 X^*_{it} + S_i(\gamma_0 + \gamma_1 X^*_{it}) \\ & + T_t[\alpha_0 + \alpha_1 X^*_{it} + S_i(\beta_0 + \beta_1 X^*_{it})] + \eta_{it} \end{aligned} \tag{1}$$

where S_i is a dummy for treated units above the cutoff, T_t is the post-period indicator and equals 1 when $T = 1$ and 0 otherwise, X^*_{it} is the normalised running variable ($X_{it} - 5,000$) and η_{it} the error component. The coefficient β_0 is the diff-in-disc estimator and identifies the effect of electing more politicians.

	Planned Budget			Actual Budget		
	Expenditures (1)	Revenues (2)	Deficit (3)	Expenditures (4)	Revenues (5)	Deficit (6)
Above 5,000 × Reform	587.4**	560.5**	-3.7	198.8	116.8	35.7
	(237.2)	(240.2)	(37.7)	(148.0)	(147.1)	(52.9)
Observations	1,355	1,428	1,198	1,440	1,668	995
R^2	0.03	0.03	0.00	0.01	0.01	0.02
Adjusted R^2	0.03	0.02	0.00	0.01	0.01	0.01

Table I.19: Diff-in-Disc Analysis with One Single Equation. Diff-in-Disc estimates. Estimation performed using WLS with triangular kernel and MSE-optimal bandwidth. No covariates incluided.

J 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 (Figure J.8), in a set of pre-treatment covariates (Figure J.9), and at most placebo cutoffs (Figure J.10).

J1 Continuity of Density

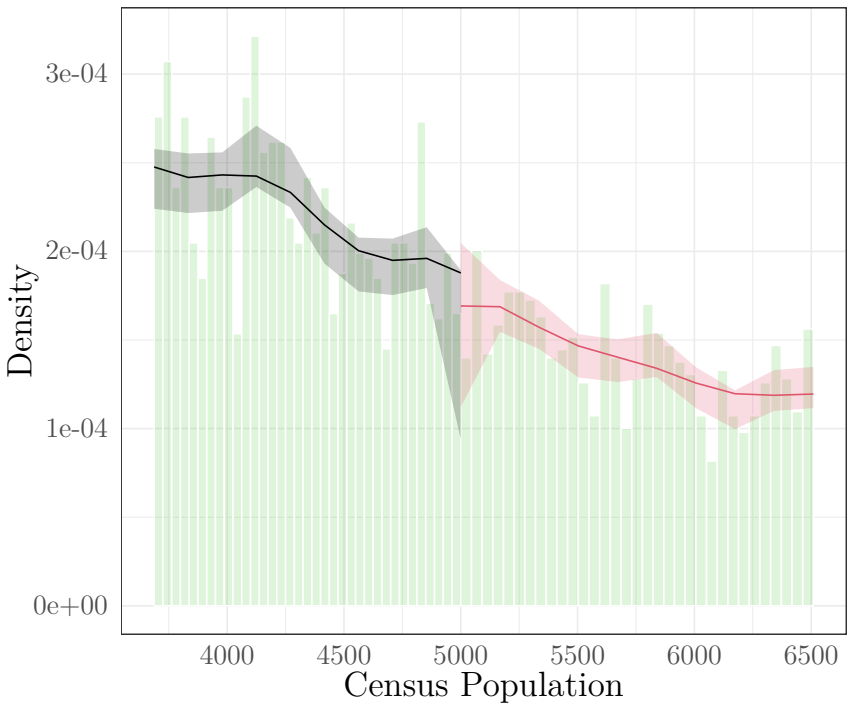

Figure J.8: Manipulation test using the local polynomial density estimators proposed by (Cattaneo et al., 2020). Histogram estimate of the running variable computed with default values in R; local polynomial density estimate (solid dark and red) and robust bias-corrected confidence intervals (shaded 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).

J2 Continuity of Potential Outcomes

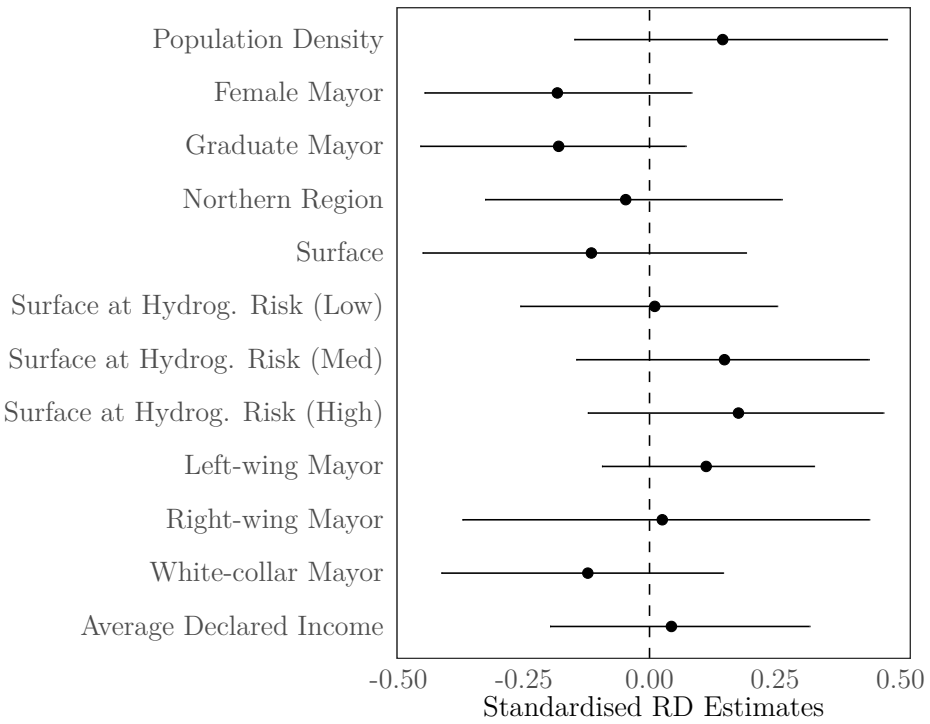

Figure J.9: Standardised RD estimates of the effect on pre-treatment covariates with 95% robust confidence intervals. Estimates constructed using local polynomial estimators with triangular kernel and CER-optimal bandwidth (as suggested by (Cattaneo et al., 2019, Ch. 5). No covariates included in the estimation. Variables used as outcomes are population density, female mayor (dummy), graduate mayor (with university degree, dummy), northern region (dummy), surface (sq.km), surface at low, medium, high hydro-geological risk (sq.km), left-wing mayor (dummy), right-wing mayor (dummy), white-collar mayor (dummy) average declared personal income of residents.

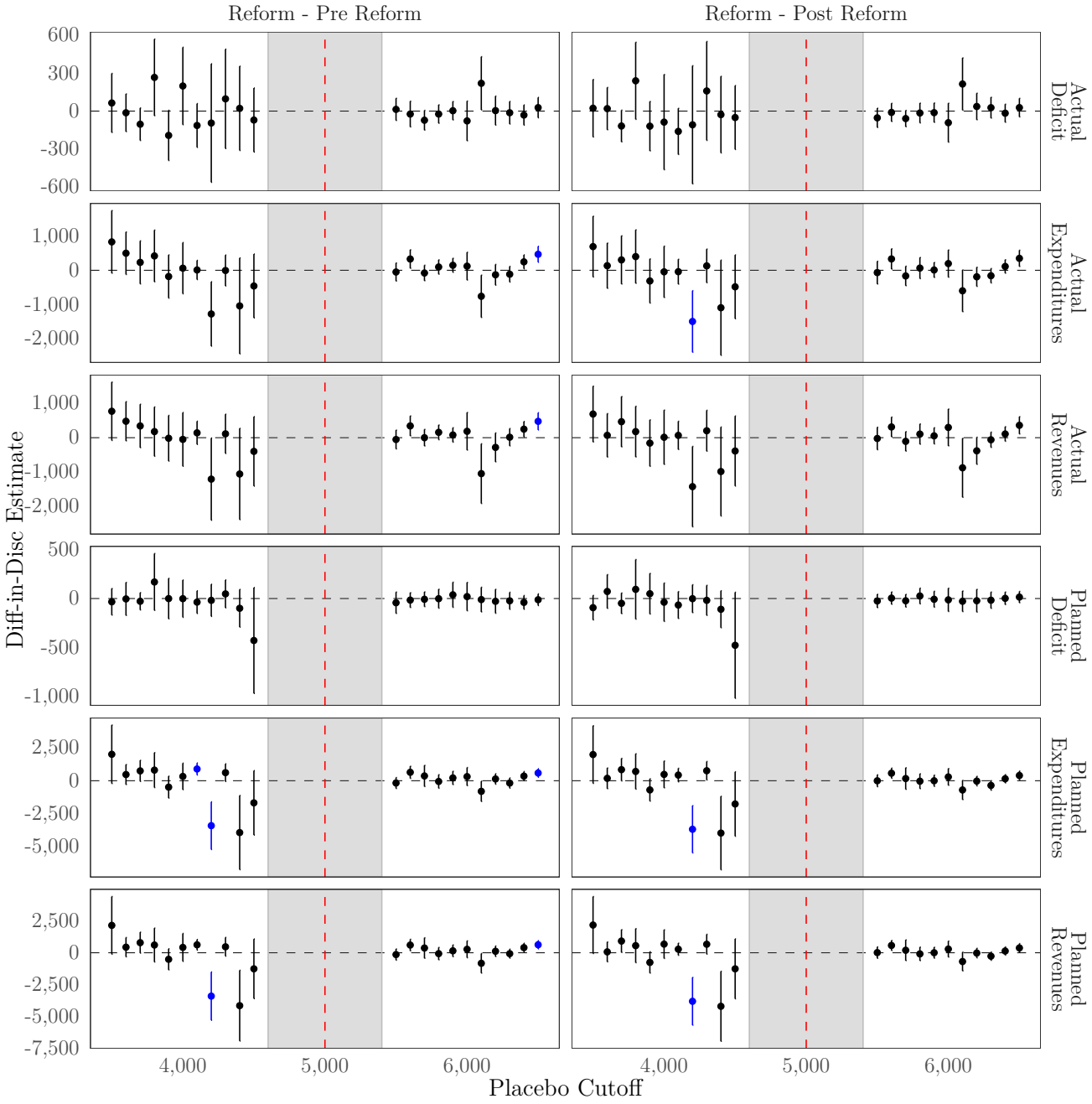

Figure J.10: Diff-in-Disc estimates with 95% robust confidence intervals. Red dashed line at the true cutoff. Blue coefficients when p.value after multiple testing adjustment smaller than 0.05. Multiple-testing adjustment performed separately for each outcome variable with Bonferroni procedure to control 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.

K 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, we compare discontinuities at the 5,000 cutoff before and after a "placebo" 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% as a result of the placebo reform. If a change in the number of politicians (both above and below the 5,000 cutoff) affected the effect of wage treatments on the outcomes, we should detect a significant difference in the discontinuity at the cutoff before and under the placebo reform. Figure K.11 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 (see Table K.20 for full regression table). A decrease in the number of politicians therefore does not change the effect of wage policies for municipalities above the 5,000 cutoff.

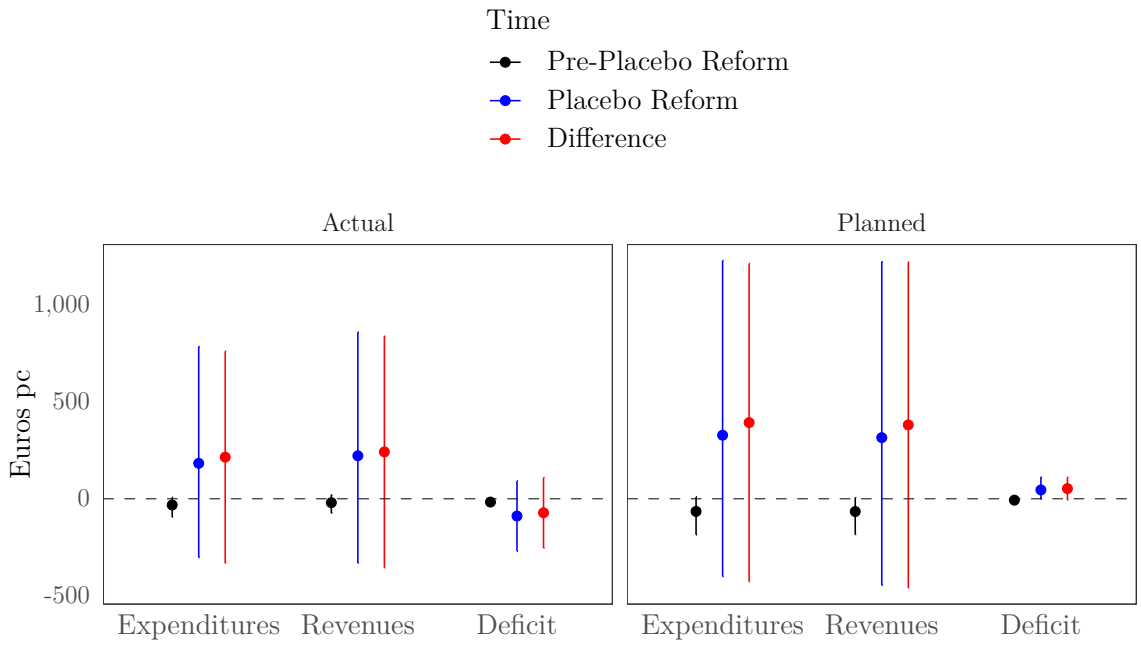

Figure K.11: Diff-in-Disc estimates with 95% robust confidence intervals estimated with alternative bandwidths. Red coefficients estimated with MSE-optimal bandwidth. Underlying RD estimates constructed using local polynomial estimators with triangular kernel and MSE-optimal bandwidth. Robust confidence interval constructed using bias correction with robust standard errors. No covariates included.

	Planned Budget			Actual Budget		
Outcome	Expenditures	Revenues	Deficit	Expenditures	Revenues	Deficit
Pre-Placebo Reform (RD)						
Estimate	-65.0	-65.8	-6.9	-31.9	-20.0	-16.8
SE	(49.9)	(48.4)	(7.4)	(25.9)	(24)	(8.4)
p.value	0.078	0.066	0.428	0.094	0.266	0.038
h	306.4	293.0	733.8	322.3	411.7	699.8
Obs. Used	2449	2301	5680	2585	3194	5412
Placebo Reform (RD)						
Estimate	328.4	315.8	45.8	183.2	222.3	-88.9
SE	(415.9)	(426.2)	(29)	(277.8)	(304)	(92.5)
p.value	0.319	0.362	0.056	0.385	0.385	0.337
h	817.7	772.8	407.4	899.1	1191.3	824.2
Obs. Used	180	164	90	216	276	180
Difference (Diff-in-Disc)						
Estimate	393.5	381.6	52.8	215.1	242.2	-72.1
SE	(418.9)	(428.9)	(30)	(279)	(304.9)	(92.9)
p.value	0.348	0.374	0.078	0.441	0.427	0.437

Table K.20: Regression table of RD results in the pre-placebo reform and placebo reform time periods as well as diff-in-disc estimates (difference in RD point estimates in the pre- and placebo reform periods) showing no statistically significant difference between the two time periods at the cutoff, suggesting that municipalities above the cutoff (paid differently) did not react differently from those below the cutoff to a same-size change in the number of politicians. No covariates included.

The "local" parallel trend assumption is indirectly tested in Figure K.12, 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 in the reform period are very similar in the pre- and post-reform periods, suggesting that, after the reform is repealed, changes at the discontinuities return to pre-reform levels.

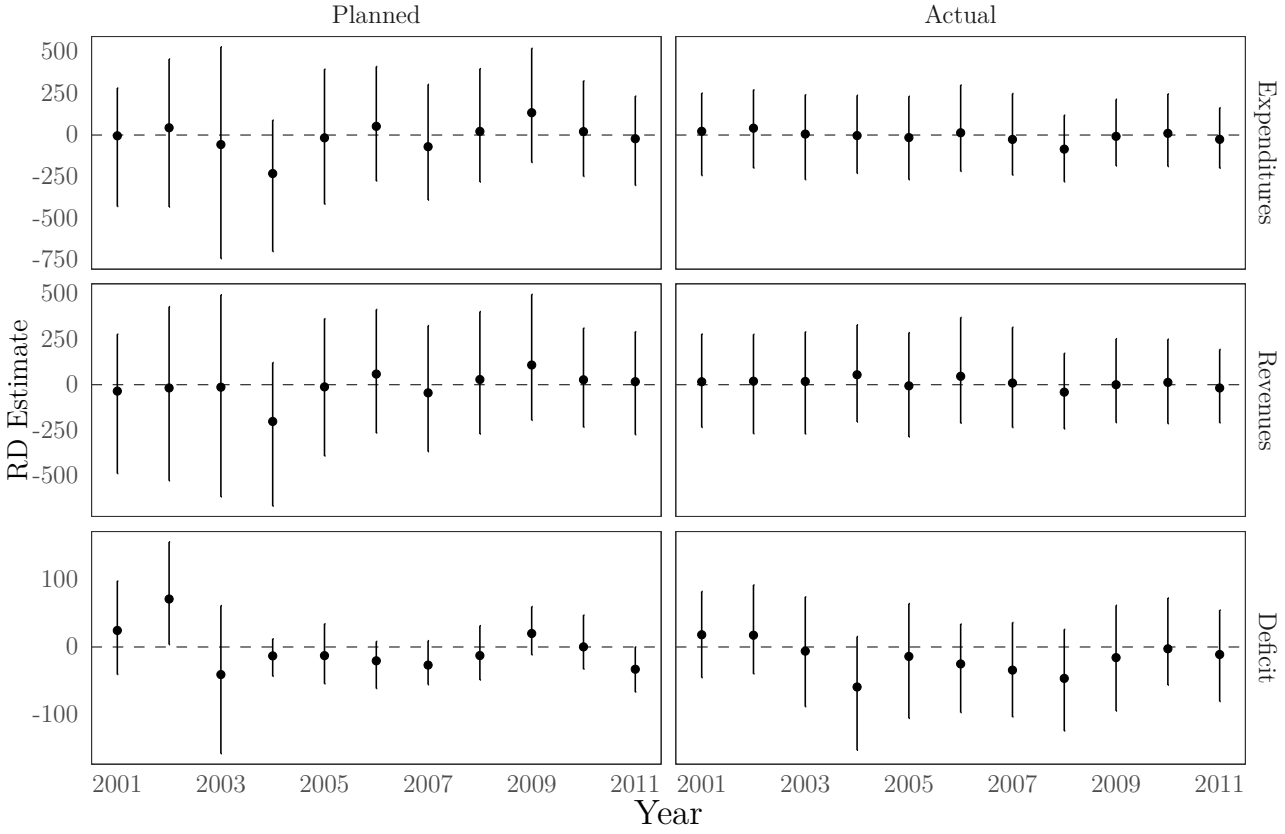

Figure K.12: RD estimates with 95% robust confidence interval for every year and every outcome in pre-treatment period (before Reform enters into force). MSE-optimal bandwidth, and triangular kernel. No covariates included.

References

Ash, E., Morelli, M. and Van Weelden, R. (2017), 'Elections and Divisiveness: Theory and Evidence', *Journal of Politics* **79**(4), 1268–1285.

Baron, D. P. and Ferejohn, J. A. (1989), 'Bargaining in Legislatures', *American Political Science Review* **83**(4), 1181–1206.

Calonico, S., Cattaneo, M. D. and Titiunik, R. (2015), 'Rdrobust: An R package for robust nonparametric inference in regression-discontinuity designs', *R Journal* **7**(1), 38–51.

Cattaneo, M. D., Idrobo, N. and Titiunik, R. (2019), *A Practical Introduction to Regression Discontinuity Designs: Foundations*, Vol. I, Cambridge University Press.

Cattaneo, M. D., Jansson, M. and Ma, X. (2020), 'Simple Local Polynomial Density Estimators', *Journal of the American Statistical Association* **115**(531), 1449–1455.

Fearon, J. D. (1999) , 'Electoral accountability and the control of politicians: Selecting good types versus sanctioning poor performance', in B. Manin, A. Prezeworski and S. Stokes, eds., '*Democracy, Accountability, and Representation*', Cambridge University Press.

Gratton, G., Guiso, L., Michelacci, C. and Morelli, M. (2021), 'From Weber to Kafka: Political instability and the Overproduction of Laws', *American Economic Review* **111**(9), 2964–3003.

Grembi, V., Nannicini, T. and Troiano, U. (2016), 'Do fiscal rules matter?', *American Economic Journal: Applied Economics* **8**(3), 1–30.

Kessler, D. and Krehbiel, K. (1996), 'Dynamics of Cosponsorship', *American Political Science Review* **90**(3), 555–566.

Klašnja, M. and Titiunik, R. (2017), 'The Incumbency Curse: Weak Parties, Term Limits, and Unfulfilled Accountability', *American Political Science Review* **111**(1), 129–148.

Mayhew, David R. (1974), *Congress: The Electoral Connection*, New Haven, CT: Yale University Press.

Morelli, M. (1999), 'Demand competition and policy compromise in legislative bargaining', *American Political Science Review* **93**(4), 809–820.

Rocca, M. S. and Gordon, S. B. (2010), 'The Position-taking Value of Bill Sponsorship in Congress', *Political Research Quarterly* **63**(2), 387–397.