Joint Discussion Paper Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 27-2023

Israel García

Budget Forecast Errors in Spanish Municipalities: The Role of Transparency

This paper can be downloaded from

https://www.uni-marburg.de/en/fb02/research-groups/economics/macroeconomics/research/magks-joint-discussion-papers-in-economics

Coordination: Bernd Hayo • Philipps-University Marburg
School of Business and Economics • Universitätsstraße 24, D-35032 Marburg
Tel: +49-6421-2823091, Fax: +49-6421-2823088, e-mail: hayo@wiwi.uni-marburg.de

Budget Forecast Errors in Spanish Municipalities: The Role of Transparency

Israel García

(University of Marburg)

This version: 1 December 2023

Corresponding author:
Israel García
Marburg Centre for Institutional Economics (MACIE)
School of Business and Economics
University of Marburg
D-35032 Marburg
Germany
Email: Israelgarcia_dominguez@hotmail.com

Budget Forecast Errors in Spanish Municipalities: The Role of Transparency

Abstract

Does budget transparency effectively limit the use of creative accounting to circumvent fiscal rules? Through examining a Spanish reform that limited the obligation to provide regular budget information, I study the effect of relaxing transparency rules on budget forecast errors. After the reform, municipalities with less than 5,000 inhabitants were not obligated to provide quarterly information on cumulative budget execution and forecast deviations. Using a difference-in-differences estimator in a sample of municipalities from this Madrid region between 2010−2019, I compare expenditure and revenue forecast errors between municipalities below 5,000 inhabitants (treatment group) and above 5,000 inhabitants (control group). I observe that reducing the frequency of budget reporting leads to a systematic underestimation of planned expenditures and revenues. Furthermore, most differences are found in pre- and electoral periods, indicating a political forecast cycle. I then combine the introduction of a gender quota with the transparency reform to study gender differences in budget forecast errors, finding that female politicians systematically deviate from initial expenditure projections for electoral purposes.

JEL Classification: C23, E62, D72, H68, J16

Keywords: Budget transparency, Fiscal rules, Difference–in–differences, Political forecast cycles, Gender, Spanish municipalities, Madrid region

1. Introduction

Since the global financial crisis, fiscal stability and sustainability have become priorities in high-income countries. Most of these countries have introduced fiscal rules to limit fiscal policy discretion and exercise tighter control over the public budget. Evidence shows that fiscal rules reduce budget deficits (Grembi et al., 2016) and constrain government debt and expenditures (Asatryan et al., 2018; Christofzik and Kessing, 2018). However, fiscal rules have also been criticised because they incentivise using fiscal gimmicks and creative accounting to circumvent them (Alesina and Perotti, 1996; Milesi-Ferretti, 2004). An important instrument is strategically using biased projections in the budgetary process (von Hagen, 2010; Picchio and Santolini, 2020).

Conceptually, budgets reflect the political determination of the citizenry, and the systematic deviation of the final budget concerning its projection might thwart the political will of the citizens (Mayper et al., 1991). A consistent over- or underestimation complicates planning; budgets become unreliable, impeding the effective steering of public finances and affecting the credibility of policy decisions (Serritzlew, 2005). Even though there is substantial evidence of budget forecast manipulation, its strategic use depends crucially on the degree of budget transparency (Milesi-Ferretti, 2004; Alt et al., 2014). Over the last two decades, transparency has become a pivotal issue in the public and political debate, given its significant and positive effects on fiscal performance. This increasing importance is reflected by the growing body of research investigating the effects of transparency on fiscal performance indicators (Alt and Lassen, 2006a; Gootjes and de Haan, 2022). However, empirical evidence of the crucial role of transparency in budget forecast errors remains remarkably scarce.

I investigate how relaxing transparency rules affects budget forecast errors by analysing the introduction of Order HAP/2082/2014 in Spain. Before this law's introduction, all municipalities were subject to the same transparency rules and regulations. After the reform, municipalities with less than 5,000 inhabitants were not obligated to provide quarterly information on cumulative budget execution or forecast deviations. I hypothesise that an exogenous change in the frequency of budget reporting induces a systematic difference in budget forecast errors. Specifically, I conjecture that relaxing reporting requirements causes an underestimation of the budget, signalling a strong commitment to fiscal restraint in the initial budget that is later overrun by larger final expenditures and revenues.

I collected data between 2010 and 2019 from the Madrid region (Comunidad Autónoma de Madrid). Municipalities in this region are subject to a homogenous set of budget rules. Furthermore, local governments must submit periodic budget information to the same territorial authority. The period 2010−2019 covers three different electoral terms and is characterised by balanced local budgets and

constraints on expenditures due to the Organic Law 2012. I estimate municipal-level regressions applying a difference-in-differences (DiD) design to compare differences in budget forecast errors before and after the reform. The control group consists of municipalities with 5,000 or more inhabitants that must report quarterly budget information. Municipalities with less than 5,000 inhabitants, only requiring annual information disclosure, comprise the treatment group.

To study the influence of transparency, I differentiate expenditures between 'mandatory spending', that is, budget items that *each municipality must provide*, and 'non-mandatory spending', services the municipality is not required to provide. I also separate revenues into 'own revenues', that is, resources under the control of local governments, and 'external revenues', unconditional transfers from supra-municipal governments. By construction, dependent variables account for size differences among municipalities, which reduces comparability concerns between the control and treatment groups.

I divide the empirical analysis into three parts. First, I compare budget forecast errors before and after the reform. Relaxing transparency requirements translates into a 350% and 200% underestimation of the average mandatory and non-mandatory forecast errors, respectively. On the revenue side, the DiD estimates show that total and external revenues are underestimated by 130% and 180%, respectively. Reducing the frequency of budget reporting leads to a systematic underestimation of the planned budget.

In the second part of the empirical analysis, I examine forecast error dynamics before and after the reform using event-study plots. The total budget must generally be balanced, which forces a similar-sized underestimation of total expenditures and revenues in 2010, 2011, and with some delay after the reform. There is also a tendency to change between non-mandatory and mandatory spending. Non-mandatory spending was significantly overestimated before the reform, while mandatory spending was underestimated only in post-reform years. On the revenue side, own resources are underestimated before and after the reform, whereas external transfers are only underestimated with some delay after the reform was introduced.

Most differences are found in pre-electoral and during electoral periods, which could be linked to a political forecast cycle. 2010−2019 is characterised by balanced local budgets and constraints on total and individual expenditures. Therefore, Political Budget Cycles (PBCs) cannot primarily work through deficits or total spending, and local politicians engage in a different strategy, namely, underestimating expenditure and revenue forecasts before an election.

Finally, in the third part of the analysis, I combine the introduction of the gender quota with the transparency reform to investigate gender differences in budget forecast errors. Focusing on

municipalities with less than 5,000 inhabitants, I use a difference-in-discontinuities estimator around the 3,000 inhabitant threshold to compare gender differences in budget deviations before and after the reform. Municipalities affected by the gender quota overestimate mandatory spending and underestimate non-mandatory spending after the reform. Furthermore, non-mandatory spending is substantially underestimated right before the election in 2015, showing that female politicians can behave as electorally motivated as male ones.

This analysis contributes to the literature on the importance of transparency for fiscal stability and sustainability. Alesina and Perotti (1996) report that a lack of transparency not only undermines fiscal rules but becomes a critical obstacle to achieving budget consolidation and expenditure control. These findings are later corroborated by the results of Alt and Lassen (2006a), Benito and Bastida (2009), Eslava (2011), and Gootjes and de Haan (2022). Similarly, Milesi-Ferretti (2004) and Alt et al. (2014) find that budget transparency is crucial for good fiscal performance and that low-transparency countries systematically circumvent fiscal rules using creative accounting. This article links transparency to one of the most important instruments used to circumvent fiscal rules: systematic deviations from budget projections. Moreover, instead of using a proxy or an index for fiscal transparency, I exploit a quasi-experimental variation in the level of transparency to provide more credible estimates.

Serritzlew (2005), Goeminne et al. (2008), Boylan (2008), Bischoff and Gohout (2010), Merola and Perez (2013), Benito et al. (2015), Boukari and Veiga (2018), and Ríos et al. (2018) have explained the existence of systematic budget forecast errors through political and socioeconomic lenses, citing government ideology and fragmentation, transparency, unemployment, economic level, and fiscal rules. However, I argue that the most relevant factor appears to be the proximity of local election dates. Overly optimistic revenue and spending forecasts tend to appear before elections, likely to increase the chances that local politicians will get re-elected. Brück and Stephan (2006) and Bohn and Veiga (2021) call this phenomenon political forecast cycles. In this article, I explore political forecast cycles using a direct measure of transparency – budget reporting frequency – to reduce endogeneity concerns.

My findings are not in line with previous work, suggesting that a higher representation of female politicians increases fiscal performance (Cabaleiro-Casal and Buch-Gómez, 2020, 2021; Balaguer-Coll and Ivanova-Toneva, 2021), reduces opportunism, and increases budget transparency (Brollo and Troiano, 2016; Araújo and Tejedo-Romero, 2016, 2018).

The remainder of this paper is organised as follows. Section 2 provides background information, Section 3 describes the data, and Section 4 describes the identification strategy. The estimation results

and robustness tests are presented in Section 5. In Section 6, I analyse gender differences, and Section 7 concludes.

2. Background information

2.1 Local government in Spain

Spanish councillors are elected every four years through a proportional representation system based on closed lists.[1] Voters express their preference for a given party by selecting the corresponding ballot. Each municipality has as many electoral ballots as parties, and each ballot includes as many candidates as the number of seats in the municipal council. The number of elected councillors is computed according to the d'Hondt law, combined with a 5% threshold to avoid the proliferation of very small parties. The order in which a party's candidates are listed determines who will be elected as councillors. Councillors choose the mayor by simple majority vote, but only candidates at the top of the respective party lists are eligible to run as mayor. There are no term limits, although, in principle, council members serve four-year terms (Organic Law 5/1985, 'General Electoral Regime'). The mayor proposes initiatives and regulations which can be passed by a majority vote from the council. The mayor controls the municipality's executive functions and explains the budget proposal to the council.[2] The council is responsible for monitoring the municipality's activities and approving the budget and any amendments.

In March 2007, the Equality Act mandated gender-balanced electoral lists requiring that at least 40% of candidates be female and at least 40% male. This quota applied to the entire party list and each section of five candidates within the list. For example, in a municipality with 13 councillors, the ballot must contain at least six women and six men, plus at least two men and two women within the first five positions and positions 6 to 10. Parties whose candidate lists do not fulfil these requirements are not allowed to participate in the elections. This quota was introduced in 2007 in municipalities with more than 5,000 inhabitants, and in 2011, it was extended to those with more than 3,000 inhabitants.

2.2 Local public finance

The Spanish public sector is divided into three levels: The State, Regional Governments, and Local Governments. Municipalities are the smallest administrative unit, but they are responsible for many tasks. The Spanish Constitution grants municipalities a notable degree of budget autonomy and flexibility, but it is also very specific about which services must be provided according to population size. All municipalities are responsible for basic services, e.g. public lighting, waste collection, and

[1] Municipalities with 250 (or less) inhabitants use an open-list system.
[2] Sweeting (2009) provides a detailed discussion of the mayor's role in Spain.

public cleaning; larger municipalities have additional obligations, such as public libraries and public transportation.

On the revenue side, municipal funds come from the Central or Regional Government and municipal resources. Municipalities levy local taxes – property tax, business tax and vehicle tax – and they collect fees and user charges. They also receive transfers from supra-municipal governments, which are mostly determined by the population.

The local government's accounting system is based on a business accounting model adapted to public sector entities. There is a duality in the accounting system; transactions are recorded when they occur and when cash settlements take place. Thus, budget reporting follows a business-like accrual-based accounting. The initial municipal budget is approved and published before the actual budget period, whereas final budget data are generally available only with considerable delay.

Since 2001, local governments must comply with the Balanced Budget Rule (BBR), stating that all planned budgets and successive modifications must generally be balanced. The Ministry of Finance has the right to veto an approved municipal budget if a violation of the BBR occurs. Local governments may incur deficits only under special circumstances and with the authorisation of the Ministry of Finance. In 2011, public budget stability was anchored in the Spanish Constitution (Article 135). A year later, the Spanish parliament approved the Organic Law 2/2012, 'Budgetary Stability and Financial Sustainability', to operationalise the budget stability obligation implemented in the Constitution. These laws prioritise debt repayment over any other expenditure and further tighten municipalities' fiscal limits by regulating government expenditures.

2.3 Transparency reforms

In October 2012, Order HAP/2105/2012 was introduced to ensure local governments complied with obligations to increase transparency, according to the guidelines in Organic Law 2/2012. The Order HAP/2105/2012 is extensive and detailed in the requirements that local governments must fulfil to abide by the principle of 'budget transparency'. It emphasises that local governments must submit periodic information to the territorial authorities in the Ministry of Finance. In the 16th article, the order establishes a quarterly obligation to provide very detailed information on expenditures and revenues. All municipalities must provide a summary of the cumulative budget execution statement at the end of each quarter, displaying all revenues and expenditures, indicating the already collected rights and obligations, and the forecast deviations.

If a local government failed to comply with the obligation to submit the required budget information, whether in terms of established deadlines or content, they would be given a non-compliant status

with the budgetary stability objective, and the territorial authorities would adopt the corrective measures provided in Organic Law 2/2012: 1) All credit and debt operations shall require the authorisation of the ministry of finance; 2) the granting of subsidies or the signing of agreements by the Central Administration will require a favourable report from the ministry of finance; 3) they will incur restrictions on budget amendments leading to a net increase in the non-financial expenditure of the local government. 4) In extreme situations where the omission of information could be considered intentional and criminal in nature, members of local government will be subject to civil and criminal liability for acts and omissions performed in the exercise of their office.

Two years later, in November 2014, Order HAP/2082/2014 was introduced to amend some of the articles previously set out by Order HAP/2105/2012. Especially relevant for my analysis is the modification of the 16th article: It now reads that local governments with a population not exceeding 5,000 inhabitants are excluded from complying with the obligation to provide budget information for the first three quarters of each year. Therefore, municipalities with less than 5,000 inhabitants only have to provide annual information on budget execution and deviations with respect to forecasts.

3. Data

3.1 Sample

The sample comprises 172 municipalities from the Madrid region (Comunidad Autónoma de Madrid), from 2010−2019. This region presents interesting data for several reasons. First, available information at the local level is generally not homogeneous between regions. Thus, building a comparable database using data from municipalities in different regions is challenging. Second, Madrid has no supra-municipal authority between municipalities and regional governments, which ensures homogeneity in terms of legal obligations with respect to public service provisions, grants, and transparency requirements.[3] At the same time, the Madrid region is diverse; it includes a *very large* municipality, some *large* ones, and a considerable number of towns and villages. Most municipalities in this sample (55%) have less than 5,000 inhabitants, and 13% have more than 50,000 inhabitants. Note that, compared to the rest of the country, this sample region underrepresents *small*

[3] The Spanish territorial organisation consists of regions (Comunidades Autónomas), provinces (Provincias), and municipalities (Municipios). Each region has one or more provinces, and provinces contain multiple municipalities. Municipalities are required to provide some mandatory services based on population. The non-mandatory services are provided by either the regional or central government. Article 36 of the local administrative law states that the provincial administration is in charge of coordinating and establishing those municipal mandatory services. According to the territorial administration, it is possible that two similarly sized municipalities that belong to the same region but are located in different provinces could have different standards of mandatory services. Madrid is a region with only one province; consequently, there cannot be any variability in municipal mandatory services.

municipalities and over-represents *large* ones.[4] This under-representation of small municipalities provides us with an almost equal number of treated and untreated municipalities.

The sample period—2010−2019—was determined by data availability. In December 2008, Order EHA/3565, 'Structure of the Budget of Local Entities', thoroughly modified the structure of local budgets. The changes affected all budgets from 2010 onwards. The budget's revenue side was virtually unchanged, but the expenditure side was significantly modified. Comparisons between most expenditure items before and after 2010 are practically impossible. The period 2010−2019 is characterised by balanced local budgets and constraints on expenditures due to Organic Law 2012. This period covers three different electoral terms from elections that took place in 2011, 2015, and 2019.

3.2 Data sources

Data for planned and actual budgets were extracted from the Ministry of Finance's CONPREL database, which I merged with data collected from municipality archives.[5] All nominal budget variables are expressed as 2010 real values. Different Spanish municipalities are required to provide different public services, which can be categorised according to population size. *Level0* encompasses very basic services which need to be performed by every municipality: public lighting, cemeteries, waste collection, public cleaning, drinking water supply, sewer system, access to urban areas, road surfacing, and food and beverage control. *Level1* contains the services municipalities with more than 5,000 inhabitants must provide: public parks, public libraries, and market and waste management. Municipalities with more than 20,000 inhabitants belong to *Level2* and provide civil defence, social services, fire safety, sports facilities, and slaughterhouses. Larger municipalities with more than 50,000 inhabitants offer *Level3* services such as public transport and environmental protection.[6] In line with this categorisation, I collect two different variables from the expenditure side, *Mandatory* spending and *Non-mandatory* spending, constructed as follows:

- *Mandatory*: For municipalities with fewer than 5,000 inhabitants, it takes the value of *Level0* expenditures. For municipalities with between 5,000 and 20,000 inhabitants, it takes the value of the sum of *Level0* and *Level1* expenditures. In municipalities with between 20,000 and 50,000 inhabitants, it takes the value of *Level0*, *Level1,* and *Level2* expenditures. For

[4] According to the Spanish Statistical Institute, in 2019, out of 8,131 municipalities, 83% had less than 5,000 inhabitants, and 5% had more than 50,000 inhabitants.

[5] https://serviciostelematicosext.hacienda.gob.es/SGFAL/CONPREL

[6] According to Law 27/2013 (27 December 2013), food and beverage control, markets, and slaughterhouses are not part of mandatory minimum services from the 2015 budget onwards. The social work category has been strongly modified. Prior to this law, municipalities were responsible for providing social services. After the reform, compulsory municipal intervention is no longer, strictly speaking, a service but a simple identification of needs. Although still a mandatory service, the extent and scope of social work conducted at the municipal level is now highly limited.

municipalities with 50,000 inhabitants or more, it takes the value of *Level0*, *Level1*, *Level2*, and *Level3* expenditures.

- *Non-mandatory:* For municipalities with fewer than 5,000 inhabitants, it takes the value of *Level1*, *Level2*, and *Level3* expenditures. For municipalities with between 5,000 and 20,000 inhabitants, it takes the value of *Level2* and *Level3* expenditures. This variable takes the value of *Level3 expenditures* in municipalities that have between 20,000 and 50,000 inhabitants.[7]

Municipal funds come from Central or Regional Government and municipal resources. According to this classification, two variables are collected from the revenue side: *Own* revenues and *External* revenues. *Own* revenues consist of those local taxes and fees under the control of local governments and whose rates, within legal limits, can be modified independently by the municipality. *External* revenues include unconditional transfers from supra-municipal governments, determined mainly by population.[8]

Socioeconomic variables were collected from the Statistical Institute of the Community of Madrid.[9] To account for the demand side of the public goods provision, I include three variables: *Unemployment rate, Share of dependents,* and *Share of immigration*, all measured in per cent. *Population* is measured in number of inhabitants, and *Rent* is a measure of municipal income, expressed in per cent of total spending.[10]

Political data were retrieved from the Ministry of the Interior's Database of Electoral Results and specific information on each municipality.[11] Given that a simple majority vote in the council can approve the budget, all that matters is the votes of the governing body (García and Hayo, 2022). I define the governing body as those councillors who are in the governing majority, either via one party's absolute majority or in a coalition.[12] *Left* measures the government's ideological orientation. When the mayor belongs to the Socialist Party or Left United, or when one of those two parties is the primary member in a coalition, the municipality is defined as left-wing oriented. *Parties in government* captures the degree of government fragmentation. *Turnout* measures the degree of voters' political involvement. *Mayor's higher education* is a dummy that takes the value 1 when the mayor has

[7] For municipalities with 50,000 inhabitants or more, all items from *Level0* to *Level3* are mandatory; therefore, they cannot be used to estimate non-mandatory spending.
[8] Financial revenues and capital transfers are excluded from the main analysis because they follow a different transparency regulation.
[9] https://gestiona.comunidad.madrid/desvan/Inicio.icm?enlace=almudena
[10] This variable is constructed based on information provided by tax authorities (as the main input), plus information about earnings, wealth, rents (capital and noncapital), social payments, and transfers in each municipality. The weight of each component is adjusted according to different factors, such as the number of households, number of declarants, age of population, percentage of rent from nonworking earnings, and a socioeconomic indicator for each municipality.
[11] https://infoelectoral.interior.gob.es/opencms/es/elecciones-celebradas/area-de-descargas/
[12] When a party has no absolute majority and a coalition is not reached, the whole council is considered the governing body.

obtained a university education. *Mayor's re-election* is a dummy that takes the value 1 when the mayor is re-elected. *Government higher education* represents the share of councillors who obtained university education. *Government high experience* is a dummy that takes the value 1 when the share of government members with previous experience in government is 50% or higher and 0 otherwise. The gender variables of interest are two dummies: *Female mayor*, taking the value 1 in case of a female mayor and 0 otherwise, and *Equality*, taking the value 1 when the share of women in government is between 40–60% and 0 otherwise.

The treatment variables of interest are two dummies and their interaction, the DiD estimate. *2015law* is a dummy that takes the value 1 for each year from 2015 onwards and 0 otherwise. Order HAP/2082/2014 was approved in November 2014, thus, the first year affected is 2015. *Small* is a dummy variable that takes the value 1 for municipalities with less than 5,000 inhabitants and 0 otherwise. Table A1 of the Appendix provides descriptive statistics.

4. Empirical strategy

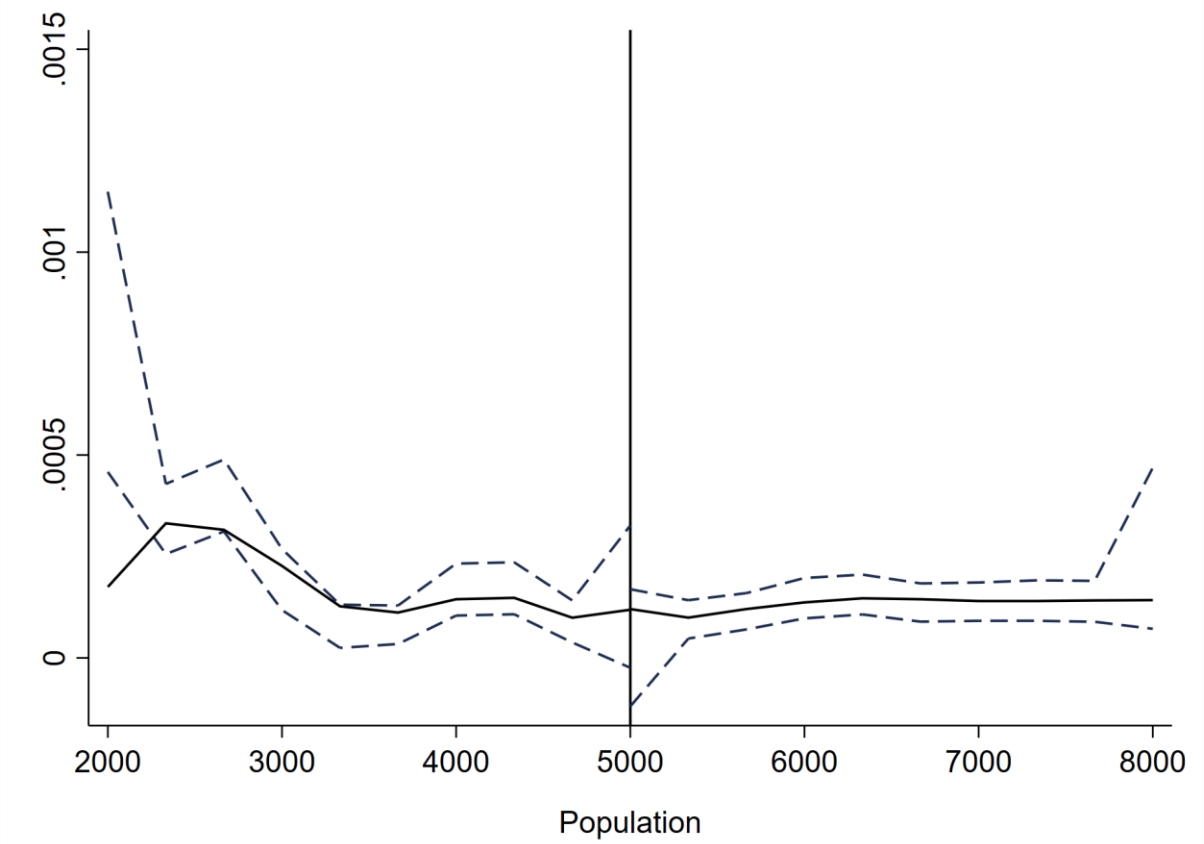
4.1 Identification

To estimate the effect of transparency rules on budget forecast errors, I exploit the introduction of Order HAP/2082/2014 and the modification introduced in its 16th Article. Before 2015, all municipalities were subject to the same transparency rules and regulations. From 2015 onwards, municipalities with less than 5,000 inhabitants are excluded from complying with the obligation to provide quarterly budget information. Municipalities with 5,000 inhabitants or more must provide a summary of the cumulative budget execution statement at the end of each quarter, indicating the already collected rights and obligations and the deviations with respect to forecasts. The control group consists of municipalities with 5,000 inhabitants or more, which must disclose budget information quarterly, whereas the treatment group consists of municipalities with less than 5,000 inhabitants, for which only annual information disclosure is required.

Before turning to the empirical model, I provide evidence that a DiD approach is suitable. In my sample of 172 municipalities, 47% (81) belong to the control group, whereas 53% (91) belong to the treatment group. A possible concern is the self-sorting of municipalities into control or treatment groups. In this case, the group assignment is based on population thresholds that are not easily manipulated. According to Foremny et al. (2017), manipulation could also be done by under- or over-reporting population statistics in a given municipality. From 2015−2019, only one municipality 'jumped' above the threshold of 5,000 inhabitants. By focusing on the 5,000 inhabitant neighbourhood, the McCrary test (2008) shows no evidence of discontinuity at the 5,000 inhabitant threshold (Figure 1). This result

is in line with Foremny et al. (2017) and Bagues and Campa (2021), who show no evidence of sorting around the threshold of 5,000 inhabitants in Spanish municipalities after 2005.

Figure 1: McCrary test – density around the 5,000 inhabitant threshold



Notes: McCrary (2008) test in densities for a second-order polynomial with evenly spaced bins. Dashed lines show 95% uniform confidence intervals.

In DiD estimation, the control group must be a good counterfactual for the treatment group to link the outcome of interest to the treatment impact and not to intrinsic differences between the control and treatment groups. Table A2 of the Appendix reports averages of the covariates in the control and treatment groups before and after the reform, as well as their differences. In the pre-treatment period, columns 1, 2, and 3, both groups differ in almost all characteristics. This picture is repeated in the post-treatment period (see columns 4, 5, and 6 of Table A2). These differences are hardly surprising since one would expect that differently-sized municipalities are dissimilar. More importantly, the fact that these differences are stable during the period analysed signifies that differences in budget forecast errors can be attributed entirely to the transparency reform and not to changes within control or treatment groups. At a 5% significance level, I discover statistical differences in *Share of dependents, Rent, Left*, *Parties in government, Turnout, and Government high experience*. Given the research question and its focus on providing information from local governments to a supra-municipal authority, the difference in the latter variable may be a reason for concern. The negative difference (column 7 of Table A3) shows a convergence between control and treatment groups in the

proportion of members of the governing body with prior experience; therefore, differences in forecast errors should be reduced through this channel. I could argue that this convergence in experience works against, rather than in favour of, the findings as it makes it more difficult to find differences between control and treatment groups post-reform.

4.2 Empirical model

I analyse the impact of transparency on budget forecast errors, measured as the difference between planned and actual budgets as a share of the total planned budget in per cent. Using a DiD estimator, I run municipal-level regressions to compare differences in budget forecast errors between municipalities with a higher frequency of budget information disclosure (control group) and municipalities allowed to disclose budget information at a lower frequency (treatment group) after the reform.

The baseline DiD estimator is of the form:

$$y_{i,t} = \varphi' X_{i,t} + \gamma Small_i + \beta(Small_i * 2015law_t) + \mu_i + \tau_t + \varepsilon_{i,t} \qquad (1)$$

where i is the municipality index, t is the year index, $y_{i,t}$ is the outcome of interest, $X_{i,t}$ is a vector of controls, μ_i is a municipality-fixed effect, τ_t is a year-fixed effect, and $\varepsilon_{i,t}$ is the idiosyncratic error term. *Small* allows us to capture the unobserved time-invariant characteristics that may differ across municipalities in the two groups. *Small*2015law* gives the DiD estimate for the reform's effect on the budget error of interest. To avoid perfect collinearity with the time-fixed effects, I omit *2015law*.

The dependent variable of interest, $y_{i,t}$ is defined as: $\frac{(Planned_{i,t} - Actual_{i,t})}{Planned\ Budget_{i,t}} * 100$; where $Planned$ refers to initial revenues and expenditures approved before the fiscal year starts. $Actual$ stands for final expenditures and revenues ascertained after the fiscal year ends. $Planned\ Budget$ indicates total initial expenditures or revenues. Thus, the dependent variable represents budget deviations as a share of the planned total budget, measured in per cent. A positive (negative) value of the variable $y_{i,t}$ indicates that the initial budget was overestimated (underestimated) by the municipality in a given year. I use three dependent variables from the expenditure side: *Mandatory, Non-mandatory,* and *Total* spending. From the revenue side, three more variables are employed: *Own*, *External,* and *Total* revenues. By using this categorisation, based on fiscal regulations, I avoid deciding, more or less arbitrarily, which expenditure or revenue items are included in the analysis. Furthermore, by construction, dependent variables account for size differences among municipalities, which reduces comparability concerns between control and treatment groups.

5. Results

5.1 Baseline

Table 1 displays the results for the baseline DiD estimation presented in Equation (1). Panel A and B show the specification without and with controls, respectively. I only report the DiD coefficient of interest, while coefficients for controls are omitted here and reported in Table A3 of the Appendix. Although the correlation between the controls and the treatment should be zero by design, Leamer (2010) argues that in the case of finite samples, correlations between the randomised treatment and the controls, by chance, tend to be non-zero, and thus, experimental randomisation of the treatment should not exclude the requirement to include further controls in the equation. Moreover, according to Hayo (2018), including control variables reduces the error of estimation and subsequently decreases the standard errors of the treatment effects. The coefficient of interest is more efficiently estimated as well.

Table 1: Baseline results

	Expenditures			Revenues		
	Total	Mandatory	Non-mandatory	Total	Own	External
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: W/o controls						
*Small*2015law*	0.44	-6.80***	-2.15***	-4.60	0.94	-2.17*
	(3.51)	(1.26)	(0.79)	(4.07)	(1.16)	(1.25)
Mean forecast error	1.02	1.67	0.78	-6.31	-1.37	-1.69
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No
Observations	1,720	1,720	1,513	1,720	1,720	1,720
Panel B: With controls						
*Small*2015law*	-2.24	-6.66***	-1.94**	-8.47**	-0.01	-2.99**
	(2.76)	(1.25)	(0.81)	(3.58)	(1.04)	(1.20)
Mean forecast error	1.02	1.67	0.78	-6.31	-1.37	-1.69
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,720	1,720	1,513	1,720	1,720	1,720

Notes: Values are based on a standard Fixed Effects estimator of Equation (1). Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

On the expenditure side shown in columns 1 to 3, there is no discrepancy between estimation with and without controls. Treatment and control municipalities do not systematically differ in their forecast errors of total spending. However, after the reform, treated (*small*) municipalities underestimated mandatory and non-mandatory spending by about 7 and 2 percentage points (pp), respectively. The average forecast error is overestimated in this sample by about 2% for mandatory spending and 1% for non-mandatory spending. Thus, there exists a reform-induced underestimation

of the budget by about 350% and 200% in absolute value for mandatory and non-mandatory spending, respectively. Furthermore, there seems to be a change from over- to underestimation of the planned budget in *small* municipalities. These findings are aligned with the empirical evidence found in Christofzik and Kessing (2018) and Repetto (2018), where an exogenous increase in budget transparency leads to lower debt and expenditure levels.[13]

By focusing on revenues, columns 4 to 6, Table 1 shows a discrepancy between the estimation without controls (Panel A) and with controls (Panel B). I consider estimation with controls a more efficient specification in this setting. Standard errors in Panel B are clearly smaller than in Panel A, which accounts for a sign of standard-error-reducing complementarity (Hayo, 2018). In addition, in Spanish local finances, the planned budget works as a legal constraint for expenditures but as a forecast for revenues. Whereas modifications on the expenditure side require a majority vote by the governing council, the budget's revenue side is treated as a mere projection of future revenue, so deviations do not require further approval. In this context, local politicians have fewer constraints on the revenue side, and municipal socioeconomic and political characteristics play a more relevant role. Total and external revenues are underestimated by about 8 and 3 pp. Given the average forecast error, total and external revenues suffer an underestimation of about 130% and 180% after the reform, respectively. Here, the treatment effect is in line with the general tendency to underestimate revenues, and it also has a lower effect in magnitude compared to the expenditure side.

In local governments, the initial municipal budget is published before the fiscal year starts, whereas final budget data are generally available only with delay and are rarely subject to external scrutiny.[14] Incumbents use budget projections to signal their future policy stance, and thus, municipal-level political debates are typically based on the initial budget, not on the final one (García and Hayo, 2021). The underestimation of the budget, reported in Table 1, suggests that politicians strategically signal a strong commitment to fiscal restraint on their initial projections that is later not implemented in the final budget. This result is in line with findings showing that voters are not pleased with high-spending governments, and incumbents must engage in alternative strategies to signal competence and/or preferences (Brender, 2003; Arvate et al., 2009; Drazen and Eslava, 2010; Eslava, 2011).

5.2 Event studies

Given that the data covers some years before and after the reform, I can also compare forecast error dynamics pre- and post-reform using event-study plots. I substitute variable *2015law* by year dummies

[13] On the contrary, Dorn et al. (2021) does not find evidence of better fiscal performance when switching to more transparent accounting standards.

[14] An exception to this statement is big infrastructure projects, for which the final costs of public provision are much higher than the planned ones; these can raise questions about poor planning and wasting public funds.

in Equation (1) and estimate the reported models again. I use 2014, the last year before the reform, as the reference year to compare annual effects pre- and post-reform.

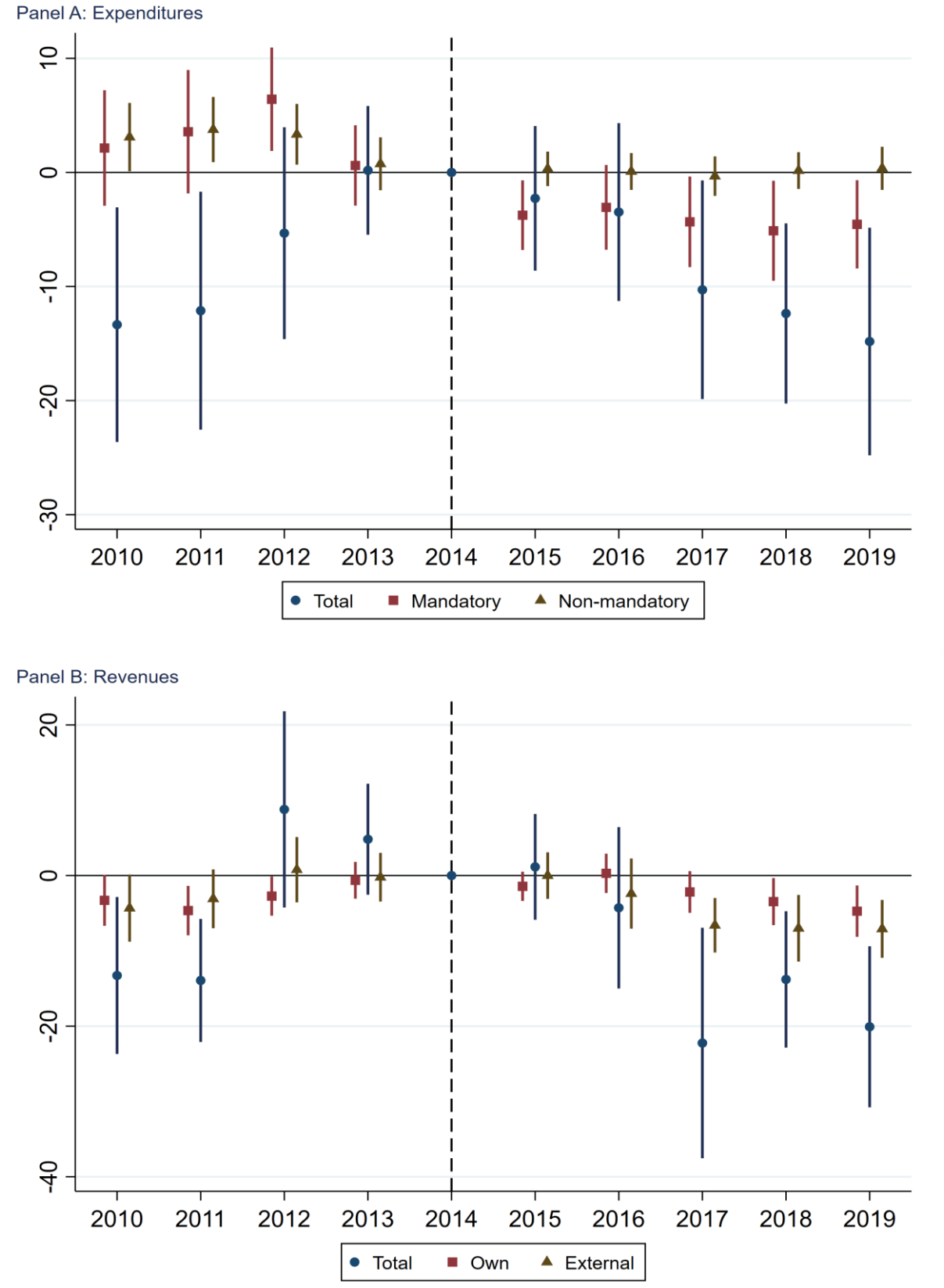
Figure 2 graphically illustrates the event-study estimates, which compare forecast error behaviours of treated versus untreated municipalities over time. Panel A shows the development of expenditure deviations, while Panel B shows revenue deviations. Total expenditure and total revenue deviations (dot markers in Panel A and B) behave very much alike before the reform, showing an underestimation of the initial budget in 2010 and 2011. Underestimation occurs with some delay after the reform, with similar-sized deviations. However, the underestimation is relatively larger for expenditures since the average forecast error is lower (see Table 1, columns 1 and 4). The total budget must generally be balanced; thus, it is unsurprising that total expenditures and revenues follow the same pattern to comply with BBRs.

Panel A shows a tendency change between mandatory and non-mandatory spending (square and triangle markers, respectively). For mandatory spending, pre-reform periods mostly show no difference in forecast errors between control and treated municipalities. However, after the reform, treated municipalities systematically underestimate mandatory expenditures by about three times the average forecast error. On the other hand, non-mandatory spending shows a very different pattern. While before the reform, treated municipalities overestimated non-mandatory spending by almost five times the average error, the entire post-reform period reports no differences in forecast errors.

Panel B shows the development of own and external forecast errors before and after the reform. Own revenues are underestimated in 2011, 2012, 2018 and 2019 by about three times the average forecast error. In contrast, pre-reform estimates show no significant differences in external revenues. After the reform, forecast errors in external revenues are underestimated in 2017, 2018 and 2019 by almost four times the average error. This underestimation is surprising since external revenues are mostly determined by population and, thus, very predictable. In line with the results obtained from expenditures, the reform seems to induce a tendency change in some budget items.

It is worth noting that, as mentioned in Section 2.3, all municipalities must provide quarterly budget information in 2013 and 2014, which makes budget deviations converge towards zero in almost every item analysed (see also Figure A1 of the Appendix). Interestingly, after the 2014 reform, there is still a two- to three-year delay until *small* municipalities start to deviate in their forecast errors. This pattern underlines the lasting influence of higher budget transparency, which hinders systematic deviations even after reporting requirements have been relaxed.

Figure 2: Baseline event-study DiD estimation



Notes: I estimate the models reported in Table 1 (Panel B), but *2015law* is substituted by year dummies. The effect is set to 0 in the last year before the reform, and the resulting reference year 2014 is marked by a dashed vertical line. Each dot, square, and triangle in the graph shows the estimated effect, and the bars indicate 95% confidence intervals.

Moreover, there is an electoral effect on budget forecast errors. The sample covers three elections (2011, 2015, and 2019), and the most significant differences are found during pre- and electoral periods. Signalling competence and/or preferences in a more or less transparent environment is the core feature of the PBC literature (Rogoff, 1990; Shi and Svensson,2006; Alt and Lassen, 2006b; Drazen and Eslava, 2010). Thus, the outcome can be linked to a different kind of electoral cycle, namely, electoral budget manipulation based on forecast errors (Brück and Stephan, 2006; Bohn and Veiga, 2021). In Spain, 2010−2019 is characterised by balanced local budgets and constraints on total and individual expenditures. In addition, Peltzman (1992) and Brender and Drazen (2008) show evidence that voters are 'fiscal conservatives'. Therefore, PBCs cannot primarily work through deficits or total spending and incumbents interested in creating PBCs must engage in a different strategy, namely, underestimating expenditures and revenues before an election. Local politicians approve of a conservative budget aligned with voters' preferences, which is then systematically overrun by higher expenditures and revenues in pre- and electoral years. This result starkly contrasts previous findings reporting overly optimistic forecasts before elections (Brück and Stephan, 2006; Boylan, 2008).

5.3 Robustness

In this subsection, I perform several robustness checks and address possible concerns related to identification. All the robustness results are reported in Table A4 in the Appendix. I start by redefining the dependent variable as budget deviations over *actual* rather than *planned* budget (Panel A). Then, I include a lagged dependent variable to account for the possible persistence of forecast errors (Panel B). Afterwards, to restrict the reform's effect to periods of high- vs low transparency, I re-estimate the model on the sample from 2013 to 2019 (Panel C). Finally, I interact the treatment with *Government high experience* to check whether results are driven by differences in experience between control and treated municipalities (Panel D).

As reported in Panel A, results are very robust when the alternative dependent variable is used. For the specifications with a lagged dependent variable (Panel B) and restricting the sample from 2013 to 2019 (Panel C), the same pattern appears. Total expenditures are now significantly underestimated, whereas non-mandatory spending is no longer significantly different between control and treatment groups post-reform. Looking at Figure 2, Panel A, it becomes clear that eliminating the years 2010 to 2012 has two effects. On the one hand, differences in total expenditures can now appear because post-reform effects are no longer diminished by forecast error underestimation from 2010 and 2011. On the other hand, non-mandatory spending becomes insignificant because pre-reform values (2010-2012) were the source of difference between control and treatment groups. Interestingly, using a lagged dependent variable as a regressor replicates the effect of restricting the analysis to high- vs low

transparency periods. On the revenue side, I do not find differences between the main specification and the robustness tests.

In Panel D, I interact the treatment with *Government high experience.* From all six categories, only in mandatory spending is the interaction of the treatment with high experience significantly negative, implying that forecast errors are underestimated in treated municipalities characterised by a high-experience government. However, the test for the equality of the two coefficients ([*Small*2015law*] and [*Small*2015law*government high experience*]) cannot reject the null hypothesis, indicating no differences according to government experience level. Only for total revenues can the equality test be rejected at the 10% level. Low-experience governments underestimate total revenues post-reform compared to high-experience governments. As commented previously, in the Spanish local budget, the planned budget works as a forecast for revenues, and differences in experience might play a more relevant role than on the expenditure side of the budget.

A more relevant issue is the comparability between control and treated municipalities. In DiD estimation, it is assumed that the control group is a good counterfactual for the treatment group, and the outcome of interest can be linked to the impact of the treatment and not to intrinsic differences between the control group and the treatment group. In this sample, the size differences among municipalities are not trivial, and thus, to alleviate concerns about comparability, I restrict the analysis to municipalities whose population is close to the 5,000 inhabitant threshold. I compare municipalities around the 5,000 inhabitant threshold before and after the transparency reform using a 'local' DiD estimator (Leuven et al., 2007; Bruce et al., 2022).

The empirical model is as follows:

$$y_{i,t} = \varphi' X_{i,t} + \gamma Small_i + \beta(Small_i * 2015law_t) + \vartheta Pop^*_{i,t} + \mu_i + \tau_t + \varepsilon_{i,t} \qquad (2)$$

where *Pop** is added to equation (1) to increase comparability between the control and treatment groups. *Pop** is the population re-centred around the 5,000 inhabitant threshold. The sample is restricted to those municipalities at a sufficiently close distance to the threshold. Table 2 reports the results at three bandwidths: 1,000, 2,500, and an 'optimal' bandwidth that tends to fall between the two, as determined by the procedure in Calonico et al. (2014). Municipalities have very similar characteristics and council size (11 or 13 councillors) in the restricted sample. Moreover, the construction of the dependent variables drastically reduces comparability concerns between control and treatment groups (Table A5 and Figure A2).

Table 2 reports the results of the 'local' DiD estimator. On the one hand, the expenditure side shows the same pattern as the baseline estimation. Mandatory and non-mandatory spending is

underestimated in treated municipalities, whereas total spending is only underestimated when the bandwidth is reduced to 1,000 inhabitants. The magnitude of effects is larger for the 'local' DiD estimation than for the baseline. In contrast to the baseline estimation, the revenue side displays a very different scenario, where neither total nor external revenues are underestimated by treated municipalities. It is important to bear in mind that population size, directly or indirectly, accounts for almost 90% of the weight in the allocation process of external revenues. Thus, almost equal-sized municipalities have little margin to significantly differ on their forecasts.

Table 2: 'Local' DiD

	Expenditures			Revenues		
	Total	Mandatory	Non-mandatory	Total	Own	External
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1,000 bandwidth						
*Small*2015law*	-11.63**	-6.94*	-4.16***	-16.80	0.04	0.88
	(4.32)	(3.40)	(0.90)	(17.33)	(3.25)	(2.60)
Mean forecast error	6.19	-0.04	0.87	-4.92	3.19	-0.86
Bandwidth	1,000	1,000	1,000	1,000	1,000	1,000
Observations	128	128	128	128	128	128
Panel B: 'Optimal' bandwidth						
*Small*2015law*	-1.63	-4.06*	-2.70***	-9.41	-0.64	0.05
	(4.33)	(2.42)	(0.82)	(7.82)	(1.49)	(2.07)
Mean forecast error	3.32	0.25	0.69	-5.09	-0.56	-1.56
Bandwidth	2094	2507	1383	2201	3242	2407
Observations	273	385	176	296	553	348
Panel C: 2,500 bandwidth						
*Small*2015law*	1.24	-4.06*	-1.76**	-5.73	-0.27	-0.76
	(3.45)	(2.42)	(0.85)	(5.79)	(2.00)	(1.86)
Mean forecast error	2.48	0.25	0.22	-4.80	-0.01	-1.49
Bandwidth	2,500	2,500	2,500	2,500	2,500	2,500
Observations	384	384	384	384	384	384

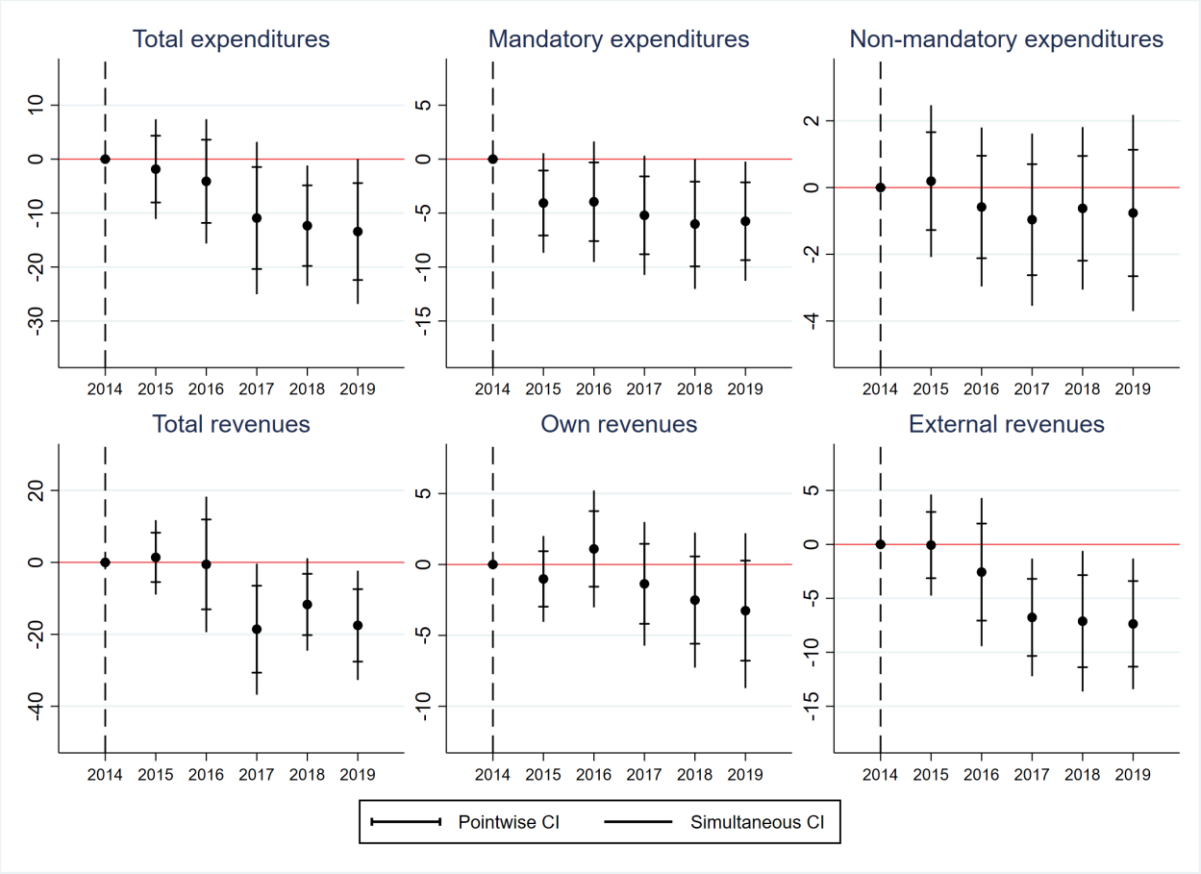
Notes: Values are based on a 'Local' Fixed Effects estimator from Equation (2). In Panel A, a bandwidth of 1,000 inhabitants is used. In Panel B, I use the 'optimal' bandwidth determined by the Calonico et al. (2014) method. In Panel C, a bandwidth of 2,500 inhabitants is used. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

In the same spirit, the parallel trend assumption can be relaxed in the context of event studies. I perform a similar analysis to the one in Bhalotra et al. (2023), where post-reform coefficients are estimated under the scenario that pre-trends might be projected forward onto the post-reform period.[15] Figure 3 shows both pointwise and simultaneous sup-t confidence bands (Olea and Plagborg-Møller, 2019). While pointwise confidence intervals permit testing preselected pointwise hypotheses,

[15] They use Rambachan and Roth (2023) "Honest DiD" estimator that accounts for estimation uncertainty, as well as uncertainty regarding the pre-treatment trend differences.

simultaneous bands are designed to include the true path of the coefficients and are therefore more convenient for giving a sense of what kinds of pre-trends are consistent with the data (Freyaldenhoven et al., 2019).

Figure 3: Post-reform event-study coefficients with simultaneous confidence intervals



Notes: I estimate the model reported in Table 1 (Panel B) in the post-reform period. *2015law* is substituted by year dummies. The effect is set to 0 in the last year before the reform, and the resulting reference year 2014 is marked by a dashed vertical line. Each post-quota coefficient from the event-study is reported along with pointwise and simultaneous 95% confidence intervals. Simultaneous sup-t confidence bands are calculated based on the code provided by Ryan Kessler: https://github.com/ryanedmundkessler/simultaneous_confidence_bands.

In line with the main results, total and mandatory spending still show significant negative deviations after the reform. However, the effect is confined to the 2019 election. On the revenue side, total and external revenues follow the same pattern found in Figure 2, showing an underestimation of the initial budget with a delay after the reform. In contrast, own revenues forecast errors are no longer different between control and treatment municipalities. Results seem to be robust, especially on the revenue side. On expenditures, the use of simultaneous confidence bands highlights the role of the electoral cycle on budget deviations. It seems plausible that the 2015 election came too soon after the reform relaxed transparency requirements to influence the behaviour of forecast errors between treated and untreated municipalities.

6. Gender differences

There is empirical evidence that women are less enthusiastic about fiscal consolidation than men (Hayo and Neumeier, 2019; García and Hayo, 2023), but at the same time, more transparent and less opportunistic (Brollo and Troiano, 2016; Araújo and Tejedo-Romero, 2016, 2018). Therefore, it is important to understand gender differences in budget forecast errors when confronted with an exogenous change in transparency rules. Since 2011, the Equality Act has mandated gender-balanced candidate lists in municipalities with more than 3,000 inhabitants. Focusing only on those *small* municipalities affected by the 2014 transparency reform, I can introduce the gender quota to investigate gender differences in budget forecast errors. I jointly exploit the budget transparency reform and the introduction of the quota to understand how relaxing transparency rules will, conditional on politicians' gender, influence budget deviations in affected municipalities. I compare *small* municipalities around the 3,000 inhabitant threshold pre- and post-reform using a 'difference-in-discontinuities' design (Casas-Arce and Saiz, 2015; Grembi et al., 2016; Gamalerio and Trombetta, 2023).

The empirical model is as follows:

$$y_{i,t} = \delta_0 + \delta_1\overline{Pop_i} + Quota_i(\gamma_0 + \gamma_1\overline{Pop_i}) + 2015law_t\left[\alpha_0 + \alpha_1\overline{Pop_i} + Quota_i(\beta_0 + \beta_1\overline{Pop_i})\right] + \varepsilon_{i,t} \quad (3)$$

where i is the municipality index, t is the year index, $y_{i,t}$ is the forecast error of interest and $\varepsilon_{i,t}$ is the idiosyncratic error term. $\overline{Pop}$ is the population re-centred around the 3,000 inhabitant threshold. *Quota* is a dummy variable that takes the value 1 when the population is above 3,000 inhabitants (0 otherwise). *Quota* allows us to capture the unobserved time-invariant characteristics that may differ across municipalities in the two groups. β_0 is the difference-in-discontinuity estimator and identifies the treatment effect of the interaction *Quota*2015law.* The main estimation is performed as in Grembi et al. (2016) or Gamalerio and Trombetta (2023), without controls and fixed effects. The sample is restricted to those municipalities below 5,000 inhabitants between 2013 and 2019. Municipalities above the 3,000 inhabitant threshold belong to the treatment group, and municipalities below comprise the control group. As shown in Figure A3 of the Appendix, there is no apparent discontinuity at the 3,000 threshold. Furthermore, no other policies were implemented based on this threshold in the relevant period (Bagues and Campa, 2021).

Table 3 shows the coefficient of interest, β_0, for an estimation with a bandwidth of 500 inhabitants, which is always close to the 'optimal' bandwidth according to the procedure proposed by Calonico et al. (2014), and yields a sufficient number of observations to estimate the models. Panel A provides the main estimation results. The impact of gender, when transparency rules are relaxed, is only relevant for mandatory and non-mandatory spending. Municipalities affected by the gender quota

overestimate mandatory spending and underestimate non-mandatory spending by about 8 pp each. Although municipalities with less than 5,000 inhabitants must provide relatively fewer mandatory services, they have small budgets and, hence, tend to focus on mandatory spending. Furthermore, standards for mandatory spending are monitored by the regional authority, which has the legal responsibility to intervene if these standards are not met. Therefore, when transparency is reduced, female politicians follow a strategy that is in line with this paper's main result, namely, to signal a strong commitment to fiscal restraint on the budget items where they have more freedom to do so. On the other hand, I find no significant differences on the revenue side, which reinforces the statement that municipalities of almost equal size have little margin to significantly differ on their revenue forecasts.

Table 3: Gender differences in forecast errors

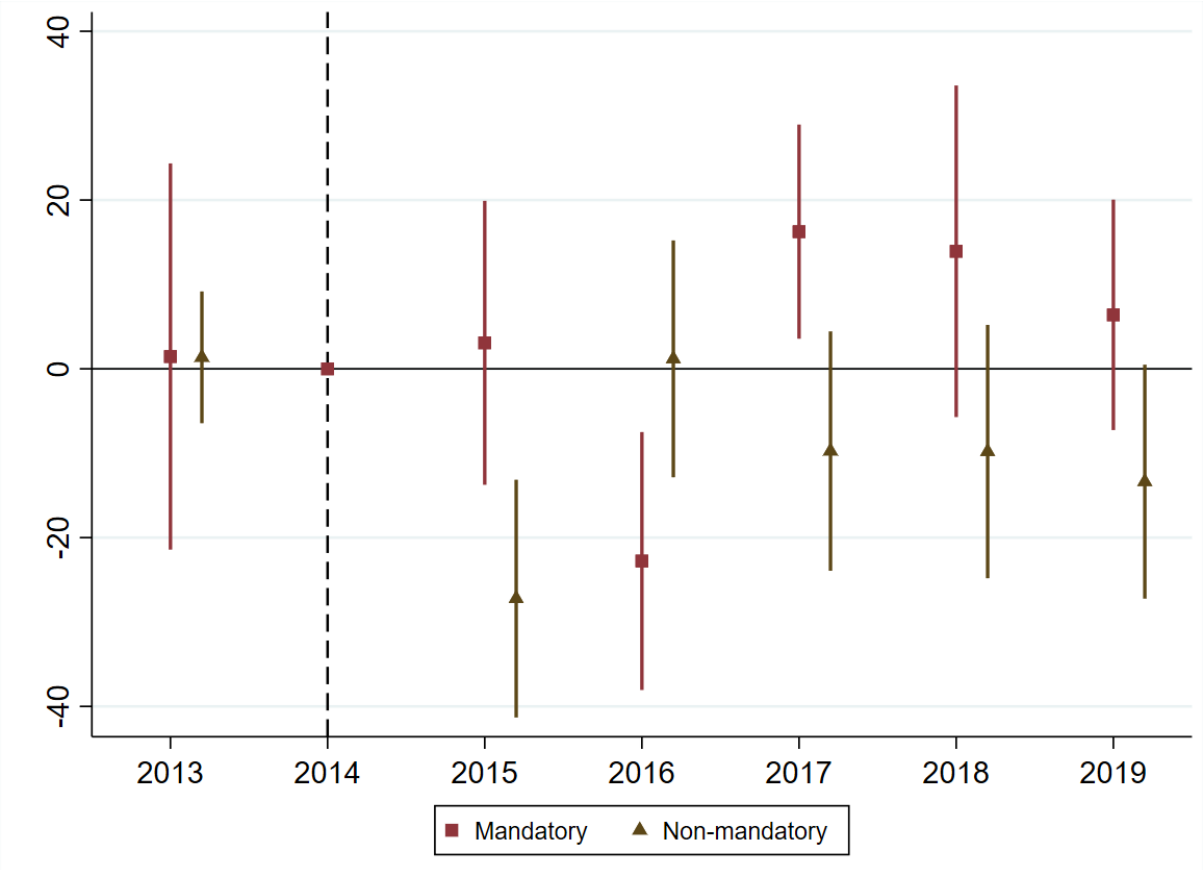
	Expenditures			Revenues		
	Total	Mandatory	Non-mandatory	Total	Own	External
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 3,000 inhabitant threshold						
*Quota*2015law*	55.90	8.47**	-8.45**	62.74	10.97	-9.84
	(51.86)	(3.20)	(3.91)	(49.15)	(7.19)	(8.20)
Mean forecast error	-1.90	-0.07	-0.53	-6.92	-2.02	-3.02
Bandwidth	500	500	500	500	500	500
Observations	82	82	82	82	82	82
Panel B: 1,900 inhabitant threshold						
*Placebo1900*2015law*	-23.27	7.69	-3.97	-19.52	9.90	-2.60
	(37.25)	(9.36)	(4.99)	(42.74)	(11.22)	(10.28)
Mean forecast error	-1.88	0.93	-0.03	-10.75	-5.85	-5.68
Bandwidth	500	500	500	500	500	500
Observations	123	123	123	123	123	123
Panel C: 4,100 inhabitant threshold						
*Placebo4100*2015law*	4.89	-2.19	-5.79	35.74	6.70	13.44
	(14.89)	(6.89)	(5.82)	(32.24)	(7.40)	(17.28)
Mean forecast error	4.73	-1.25	-0.30	-0.31	2.00	-1.19
Bandwidth	500	500	500	500	500	500
Observations	53	53	53	53	53	53

Notes: Values are based on a 'difference-in-discontinuities' estimator from Equation (3). Panel A reports the results of relaxing transparency rules in municipalities below/above the 'true' *Quota* threshold (3,000). Panel B and C report the results of relaxing transparency rules in municipalities below/above artificial *Quota* thresholds (1,900 and 4,100). Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

It is interesting to see that differences in forecast errors are driven by a higher share of female councillors in the governing body and not by female mayors. Looking at Figure A4, it is clear that there is no apparent discontinuity in the probability of having a female mayor at the 3,000 inhabitant

threshold (Panel A). However, there is an increase in the likelihood of having a gender-balanced government right below and above the threshold (Panel B). Municipalities affected by the gender quota are more likely to have gender-balanced governments, which also affects forecast errors. This feature is aligned with evidence found by García and Hayo (2022, 2023), which shows that the share of women in government is very important at the local level, more relevant even than having a female mayor. Therefore, focusing only on the mayor, the standard approach in the literature, likely leads to an incomplete picture of gender differences in municipal-level policymaking.

Figure 4: 'Difference-in-discontinuities' event-study



Notes: I estimate the models reported in Table 3 (Panel A), but *2015law* is substituted by year dummies. The effect is set to 0 in the last year before the reform, and the resulting reference year 2014 is marked by a dashed vertical line. Each square and triangle in the graph shows the estimated effect, and the bars indicate 90% confidence intervals.

I performed a placebo test to evaluate whether the results could arise by chance. I estimate the same model with the same bandwidth but at false population thresholds. I chose a 1,900 and 4,100 inhabitant threshold so the bandwidths would not overlap with those of the 'correct' quota threshold. Results can be found in Panel B and C of Table 3. At these false thresholds, there is no difference in forecast errors between treated and untreated municipalities. It should also be noted that the main effect found in Panel A is sensitive to the chosen bandwidth, which is not uncommon in this type of analysis (Grembi et al., 2016). In this study, mandatory spending is more robust than non-mandatory spending to a bandwidth increase, but statistical significance is lost at a 1,000 inhabitant bandwidth

for both. This is not a surprise since the probability of having a gender-balanced government becomes very similar for quota and non-quota municipalities, when the distance to the 3,000 inhabitant threshold increases (Figure A4, Panel B).

I analyse forecast error dynamics using an event-study plot for mandatory and non-mandatory spending in the context of a 'difference-in-discontinuities' design. Estimates are illustrated graphically in Figure 4. As discussed before, mandatory spending (square markers) does not seem to support the political forecast cycle theory. However, since mandatory spending is the main focus of *small* municipalities, it could be possible that the room to manoeuvre is constrained on pre- and electoral years, and differences can only appear in non-electoral years (2016 and 2017). On the contrary, non-mandatory spending (triangle markers) is substantially underestimated right before the election in 2015. Applying a mixed-gender close election analysis, García and Hayo (2022) find similar behaviour in female politicians who cut non-mandatory spending in electoral years. This shows that female politicians can behave as opportunistically as males and, under specific circumstances, they have to signal their competence and/or preferences much more strongly than their male colleagues.

To further address doubts about the robustness of the results, I replicate the analysis in Table 3 and Figure 4 using a 'local' DiD estimator, which allows to control for municipality- and year-fixed effects (Leuven et al., 2007; Bruce et al., 2022). As reported in Table A6, results remain unchanged when using the alternative estimator and a rich set of fixed effects. Furthermore, non-mandatory spending in Figure A5 is now substantially underestimated during electoral periods (2015 and 2019) reinforcing the opportunistic behaviour of female politicians.

7. Conclusion

A widespread assumption in the literature is that fiscal rules incentivise creative accounting to circumvent them, especially in environments with low budget transparency (Alesina and Perotti, 1996; Milesi-Ferretti, 2004; Alt et al., 2014). In light of this, I use a dataset on Spanish municipalities to investigate the effects of relaxing transparency rules on budget forecast errors. The strategic use of biased projections in the budgetary process is a very important instrument for circumventing fiscal rules (von Hagen, 2010; Picchio and Santolini, 2020). However, empirical evidence on the crucial role of transparency in budget forecast errors is remarkably scarce (Ríos et al., 2018). The general hypothesis of this study is that an exogenous change in the frequency of budgetary information reporting will induce a systematic difference in budget forecast errors. Specifically, I conjecture that relaxing reporting requirements causes an underestimation of the budget, signalling a strong

commitment to fiscal restraint in the initial budget that is later overrun by larger final expenditures and revenues.

In November 2014, Order HAP/2082/2014 was introduced in Spain. Before the law's introduction, all municipalities were subject to the same transparency rules and regulations. After the reform, municipalities with less than 5,000 inhabitants were not obligated to provide quarterly information on the cumulative budget execution and deviations with respect to forecasts. Using a DiD estimator, I compare differences in budget forecast errors pre- and post-reform between municipalities with 5,000 or more inhabitants (control group) and below 5,000 inhabitants (treatment group).

First, I compare budget forecast errors before and after the reform. I find that relaxing transparency requirements translates into a 350% and 200% underestimation of the average mandatory and non-mandatory forecast error, respectively. On the revenue side, the DiD estimates show that total and external revenues are underestimated by 130% and 180%, respectively.

Second, total expenditures and revenues follow the same annual pattern to comply with fiscal rules. There is also a tendency change between non-mandatory and mandatory spending, where only mandatory spending is underestimated in post-reform years. On the revenue side, own resources are underestimated before and after the reform, whereas external transfers are only underestimated with some delay after the reform was introduced. Findings are mostly present in pre- and electoral periods, linking the results to a political forecast cycle.

Third, I jointly exploit the introduction of the gender quota with the transparency reform to investigate gender differences in budget forecast errors when transparency rules are relaxed. Focusing on municipalities with less than 5,000 inhabitants, I find that municipalities affected by the gender quota overestimate mandatory spending and underestimate non-mandatory spending. Furthermore, non-mandatory spending is substantially underestimated right before the election in 2015, showing that female politicians must signal their competence and/or preferences much more strongly than their male colleagues.

Corroborating my initial conjecture, this analysis presents evidence that an exogenous reduction of transparency rules causes a systematic underestimation of budget projections. This result could be explained in light of previous evidence that suggests voters are not pleased with high-spending governments. Therefore, when transparency is reduced, incumbent politicians can signal a strong commitment to fiscal restraint that is later not implemented in the final budget. In addition, the results challenge the established opinion that incumbents approve over-optimistic initial budgets to signal competence and/or preferences before elections. It instead shows that initial expenditures and

revenues tend to be underestimated to please 'fiscal conservative' voters when an election is approaching.

It is worth noting that compared to the rest of the country, the sample region under-represents small municipalities and over-represents large ones, which might endanger the external validity of the analysis. However, according to Profeta and Woodhouse (2022, 22): 'external validity is ultimately best addressed by comparing the results of several internally valid studies conducted in different contexts and at different points in time'. This study offers high internal validity due to the highly homogenous set of budget and transparency rules and fulfilling the requirements for valid difference-in-differences estimation.

To summarise, in contrast to studies reporting the use of overly optimistic revenue and spending forecasts to circumvent fiscal rules, especially before elections (Brück and Stephan, 2006; Bohn and Veiga, 2021), I show that reducing budget transparency leads to a systematic underestimation of initial revenue and expenditures. This result is more in line with findings showing that voters are not pleased with high-spending governments (Brender, 2003; Drazen and Eslava, 2010). I also provide evidence that incumbents engage in political forecast cycles to signal competence and/or preferences in pre- and electoral periods. In addition, female politicians appear to use forecast errors opportunistically once transparency rules are relaxed.

References

Alesina, A. and Perotti, R. (1996). Fiscal discipline and the budget process. *American Economic Review* 86, 401–407.

Alt, J.E. and Lassen, D.D. (2006a). Transparency, political polarization, and political budget cycles in OECD countries. *American Journal of Political Science* 50, 530–550.

Alt, J.E. and Lassen, D.D. (2006b). Fiscal transparency, political parties, and debt in OECD countries. *European Economic Review* 50, 1403–1439.

Alt, J.E., Lassen, D.D., and Wehner, J. (2014). It isn't just about Greece: domestic politics, transparency, and fiscal gimmickry. *British Journal of Political Science* 44, 707–716.

Araujo, J. F. F. E. and Tejedo-Romero, F. (2016). Women's political representation and transparency in local governance. *Local Government Studies* 42, 885–906.

Araujo, J. F. F. E. and Tejedo-Romero, F. (2018). Does gender equality affect municipal transparency: The case of Spain. *Public Performance and Management Review* 41, 69–99.

Arvate, P.R., Avelino, G., and Tavares, J. (2009). Fiscal conservatism in a new democracy: "Sophisticated" versus "naïve" voters. *Economics Letters* 102, 125–127.

Asatryan, Z., Castellón, C., and Stratmann, T. (2018). Balanced budget rules and fiscal outcomes: Evidence from historical constitutions. *Journal of Public Economics* 167, 105–119.

Bagues, M. and Campa, P. (2021). Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design. *Journal of Public Economics* 194, 104315.

Balaguer-Coll, M.T. and Ivanova-Toneva, M. (2021). The impact of women's leadership in local government: The case of Spain. *International Public Management Journal* 24, 455–475.

Benito, B, and Bastida, F. (2009). Budget transparency, fiscal performance, and political turnout: An international approach. *Public Administration Review* 69, 403–417.

Benito, B, Guillamón, M. D., and Bastida, F. (2015). Budget forecast deviations in municipal governments: Determinants and implications. *Australian Accounting Review* 72, 45–70.

Bhalotra, S., Clarke, D., Gomes, J. F., and Venkataramani, A. (2023). Maternal mortality and women's political power. *Journal of the European Economic Association* 21, 2172–2208.

Bischoff, I. and Gohout, W. (2010). The political economy of tax projections. *International Tax and Public Finance* 17, 133–150.

Bohn, F. and Veiga, F. J. (2021). Political forecast cycles. *European Journal of Political Economy* 66, 101934.

Boukari, M. and Veiga, F. (2018). Disentangling political and institutional determinants of budget forecast errors: A comparative approach. *Journal of Comparative Economics* 46, 1030–1045.

Boylan, R. T. (2008). Political distortion in state forecast. *Public Choice* 136, 1091–1212.

Brender, A. (2003). The effect of fiscal performance on local government election results in Israel: 1989–1998. *Journal of Public Economics* 87, 2187–2205.

Brender, A. and Drazen, A. (2008). How do budget deficits and economic growth affect re-election prospects? Evidence from a large panel of countries. *American Economic Review* 98, 2203–2220.

Brollo, F. and Troiano, U. (2016). What happens when a woman wins an election? Evidence from close races in Brazil. *Journal of Development Economics* 122, 28–45.

Brück, T. and Stephan, A. (2006). Do Eurozone countries cheat with their budget forecasts? *Kyklos* 59, 3–15.

Cabaleiro-Casal, R. and Buch-Gómez, E. (2020). Women in Spanish municipal councils and budgetary policies. *Urban Affairs Review* 56, 1715–1745.

Cabaleiro-Casal, R. and Buch-Gómez, E. (2021). Female politicians in municipal councils and fiscal performance. *Economics & Politics* 33, 289–314.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust non-parametric confidence intervals for regression-discontinuity designs. *Econometrica* 82, 2295–2326.

Casas-Arce, P. and Saiz, A. (2015). Women and power: Unpopular, unwilling, or held back? *Journal of Political Economy* 123, 641–669.

Christofzik, D.I. and Kessing, S.G. (2018). Does fiscal oversight matter? *Journal of Urban Economics* 105, 70–87.

Dorn, F., Gaebler, S., and Roesel, F. (2021). Ineffective fiscal rules? The effect of public sector accounting standards on budgets, efficiency, and accountability. *Public Choice* 186, 387–412.

Drazen, A. and Eslava, M. (2010). Electoral manipulation via voter-friendly spending: Theory and evidence. *Journal of Development Economics* 92, 39–52.

Eslava, M. (2011). The political economy of fiscal deficits: A survey. *Journal of Economic Surveys* 25, 645–673.

Foremny, D., Jofre-Monseny, J., and Solé-Ollé, A. (2017). 'Ghost citizens': Using notches to identify manipulation of population-based grants. *Journal of Public Economics* 154, 49–66.

Freyaldenhoven, S., Hansen, C., and Shapiro, J. (2019). Pre-event trends in the panel event-study design. *American Economic Review* 109, 3307–3338.

Gamalerio, M. and Trombetta, F. (2023). Fiscal rules and the selection of politicians: theory and evidence from Italy. Available at SSRN: http://dx.doi.org/10.2139/ssrn.3035437

García, I. and Hayo, B. (2021). Political budget cycles revisited: Testing the signalling process. *European Journal of Political Economy* 69, 102030.

García, I. and Hayo, B. (2022). The influence of politicians' sex on political budget cycles: An empirical analysis of Spanish municipalities. *MAGKS Discussion Paper Series* 23–2022.

García, I. and Hayo, B. (2023). Fiscal reform in Spanish municipalities: Gender differences in budgetary adjustment. *MAGKS Discussion Paper Series* 06–2023.

Goeminne, S., Geys, B., and Smolders, C. (2008). Political fragmentation and projected tax revenues: Evidence from Flemish municipalities. *International Tax and Public Finance* 15, 297–315.

Gootjes, B. and de Haan, J. (2022). Do fiscal rules need budget transparency to be effective? *European Journal of Political Economy* 75, 102210.

Grembi, V., Nannicini, T., and Troiano, U. (2016). Do Fiscal Rules Matter? *American Economic Journal: Applied Economics 8*, 1–30.

Hayo, B. (2018). On standard-error-decreasing complementarity: Why collinearity is not the whole story. *Journal of Quantitative Economics* 16, 289–307.

Hayo, B. and Neumeier, F. (2019). Public preferences for government spending priorities: Survey evidence from Germany. *German Economic Review* 20, e1–e37, 2019.

Leamer, E. E. (2010). Tantalus on the road to Asymptopia. *Journal of Economic Perspectives* 24, p. 31–46.

Mayper, A. G., Granof, M., and Giroux, G. (1991). An analysis of municipal budget variances. *Accounting, Auditing and Accountability Journal* 4, 29–50.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: a density test. *Journal of Econometrics* 142, 698–714.

Merola, R. and Pérez, J. J. (2013). Fiscal forecast errors: Governments vs independent agencies? *European Journal of Political Economy* 32, 285–299.

Milesi-Ferretti, G. M. (2004). Good, bad or ugly? On the effects of fiscal rules with creative accounting. *Journal of Public Economics* 88, 377–394.

Olea, J. L. M. and Plagborg-Møller, M. (2019). Simultaneous confidence bands: Theory, implementation and an application to SVARs. *Journal of Applied Econometrics* 34, 1–17.

Peltzman, S. (1992). Voters as fiscal conservatives. *Quarterly Journal of Economics* 107, 327–361.

Picchio, M. and Santolini, R. (2020). Fiscal rules and budget forecast errors of Italian municipalities. *European Journal of Political Economy* 64, 101921.

Profeta, P. and Woodhouse, E. F. (2022). Electoral rules, women's representation and the qualification of politicians. *Comparative Political Studies* 55, 1471–1500.

Rambachan, A. and Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies* 90, 2555–2591.

Repetto, L. (2018). Political budget cycles with informed voters: Evidence from Italy. *Economic Journal* 128, 3320–3353.

Ríos, A. M., Guillamón, M. D., Benito, B., and Bastida, B. (2018). The influence of transparency on budget forecast deviations in municipal governments. *Journal of Forecasting* 37, 457–474.

Rogoff, K. (1990). Equilibrium political budget cycles. *American Economic Review* 80, 21–36.

Serritzlew, S. (2005). Breaking budgets: An empirical examination of Danish municipalities. *Financial Accountability and Management* 21, 413–435.

Shi, M. and Svensson, J. (2006). Political budget cycles: Do they differ across countries and why? *Journal of Public Economics* 90, 1367–1389.

von Hagen, J. (2010). Sticking to fiscal plans: The role of institutions. *Public Choice* 144, 487–503.

Appendix .....

Table A1: Descriptive statistics

Variable	*Description*	*Min*	*Max*	*Mean*	*Std. Dev.*
Total expenditures	Forecast error in total expenditures in relation to planned total expenditures (in %).	-414.95	76.93	1.02	29.24
Mandatory expenditures	Forecast error in mandatory expenditures in relation to planned total expenditures (in %).	-104.87	64.6	1.67	10.95
Non-mandatory expenditures	Forecast error in non-mandatory expenditures in relation to planned total expenditures (in %).	-48.37	73.22	0.78	6.17
Total revenues	Forecast error in total revenues in relation to planned total revenues (in %).	-650.6	81.38	-6.31	36.02
Own revenues	Forecast error in own revenues in relation to planned total revenues (in %).	-86.04	61.02	-1.37	9.71
External revenues	Forecast error in external revenues in relation to planned total revenues (in %).	-101.03	77.01	-1.69	13.72
Population	Number of inhabitants.	44	3273049	37775.28	246141.42
Unemployment rate	Percentage of registered unemployed in relation to labour force (in %).	0	85.51	32.47	17.29
Share of dependents	Share of the population below 15 and over 65 in relation to total population (in %).	18.18	57.38	32.11	3.84
Share of immigration	Share of non-Spanish inhabitants in relation to total population (in %).	0	40.42	13.17	5.62
Rent	Municipal income in relation to total expenditures (in %).	73.98	3600.17	1382.49	621.69
Left	Dummy variable taking value 1 when the governing body has left-wing ideology (0 otherwise).	0	1	.22	.42
Parties in government	Number of parties in the government.	1	7	1.82	1.31
Turnout	Number of votes in relation to total possible votes in local elections (in %).	39.17	100	74	7.58
Mayor's higher education	Dummy variable taking value 1 when the mayor obtained university education (0 otherwise).	0	1	.39	.49
Mayor's re-election	dummy variable taking the value 1 when the mayor is re-elected	0	1	.53	.5
Government higher education	Share of the governing body with university education (in %).	0	100	28.02	26.56
Government high experience	dummy variable taking the value 1 when the share of government members with previous experience in government is 50% or higher (0 otherwise).	0	1	.44	.5

Female mayor	Dummy variable taking value 1 when mayor is female (0 otherwise).	0	1	.24	.42
Equality	Dummy variable taking the value 1 when the share of women in government is between 40–60% (0 otherwise).	0	1	.56	.5
2015 law	Dummy variable taking the value 1 from 2015 onwards (0 otherwise).	0	1	.5	.5
Small	Dummy variable that takes the value 1 for municipalities with less than 5,000 inhabitants (0 otherwise).	0	1	.53	.5

Table A2: Covariates before and after the 2015 transparency reform

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Pre-			Post-		DiD
	Treatment	Control	Difference	Treatment	Control	Difference	(3)-(6)
Population	1508.12	79551.98	-78043.87***	1509.31	79219.07	-77709.47***	-334.11
	(60.42)	(18016.32)	(16720.49)	(60.99)	(17613.99)	(16539.47)	(23518.6)
Unemployment rate	39.99	34.47	5.52***	29.92	24.76	5.16***	0.37
	(0.94)	(0.65)	(1.18)	(0.85)	(0.58)	(1.05)	(1.58)
Share of dependents	32.98	29.68	3.29***	33.96	31.40	2.56***	0.73**
	(0.20)	(0.13)	(0.25)	(0.19)	(0.11)	(0.23)	(0.34)
Share of immigration	14.49	15.02	-0.52	11.15	12.13	-0.98**	0.46
	(0.32)	(0.23)	(0.41)	(0.26)	(0.18)	(0.33)	(0.52)
Rent	974.66	1630.18	-665.53***	1125.35	1897.00	-771.65***	116.12**
	(22.58)	(22.07)	(31.80)	(26.60)	(23.99)	(36.19)	(48.18)
Left	0.20	0.16	0.04	0.23	0.31	-0.08**	0.12**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
Parties in government	1.35	1.78	-0.43***	1.45	2.81	-1.36***	0.94***
	(0.03)	(0.06)	(0.07)	(0.04)	(0.09)	(0.09)	(0.11)
Turnout	79.89	69.62	10.27***	76.85	68.33	8.52***	1.74**
	(0.33)	(0.20)	(0.40)	(0.34)	(0.18)	(0.40)	(0.56)
Mayor's higher education	0.25	0.54	-0.29***	0.23	0.57	-0.34***	0.05
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
Mayor's re-election	0.59	0.51	0.08**	0.57	0.45	0.12***	-0.04
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)	(0.05)
Government higher education	17.97	41.43	-23.46***	16.34	39.53	-23.19***	-0.27
	(1.04)	(1.32)	(1.66)	(0.99)	(1.28)	(1.60)	(2.31)
Government high experience	0.35	0.55	-0.20***	0.40	0.46	-0.06*	-0.13**
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.48)
Female mayor	0.20	0.26	-0.06**	0.23	0.25	-0.02	-0.04
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.04)
Equality	0.39	0.64	-0.25***	0.50	0.75	-0.25***	0.00
	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.03)	(0.05)

Notes: Difference indicates a t-test for equal means across the given treatment dimensions. *, **, and *** indicate significance at a 10%, 5%, and 1% level, respectively.

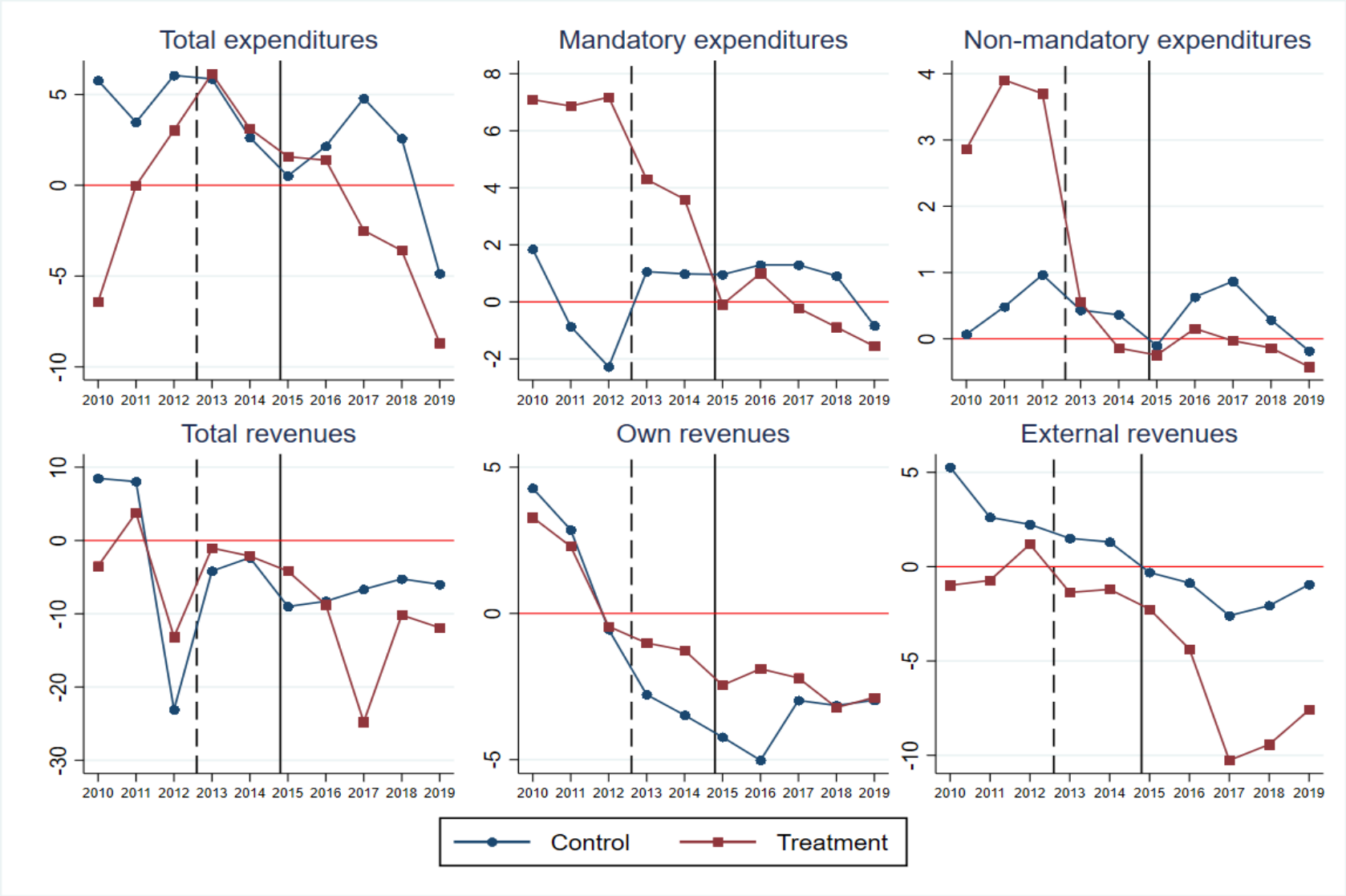
Table A3: Baseline results with full set of covariates

	Expenditures			Revenues		
	Total	Mandatory	Non-mandatory	Total	Own	External
	(1)	(2)	(3)	(4)	(5)	(6)
Population	0.00**	-0.00*	0.00	0.00**	0.00	0.00
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Unemployment rate	0.08	0.01	-0.03*	0.01	0.03	0.04
	(0.11)	(0.05)	(0.02)	(0.12)	(0.03)	(0.05)
Share of dependents	1.15	0.73**	-0.05	1.43**	-0.24	0.97***
	(0.87)	(0.37)	(0.17)	(0.64)	(0.17)	(0.32)
Share of immigration	1.42	0.55**	-0.02	0.82	0.06	0.20
	(0.95)	(0.25)	(0.14)	(0.61)	(0.19)	(0.29)
Rent	-0.05***	-0.01***	-0.00	-0.06***	-0.01***	-0.01***
	(0.01)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
Left	-8.12**	-0.89	0.20	-13.29***	-1.68*	-1.33
	(3.99)	(1.14)	(1.00)	(5.00)	(0.88)	(0.95)
Parties in government	1.13	0.10	0.04	2.16**	0.66**	0.09
	(0.69)	(0.30)	(0.18)	(0.91)	(0.27)	(0.34)
Turnout	0.66	0.12	0.08	0.33	0.03	-0.13
	(0.45)	(0.17)	(0.09)	(0.34)	(0.07)	(0.12)
Mayor's higher education	5.61*	1.12	0.30	7.05*	0.43	2.69**
	(3.11)	(1.00)	(0.86)	(3.76)	(0.87)	(1.35)
Mayor's re-election	1.70	-1.37	-0.21	2.63	1.40**	-0.06
	(1.32)	(0.83)	(0.40)	(2.12)	(0.54)	(0.76)
Government higher education	-0.02	-0.02	-0.02	-0.06	-0.02	-0.03
	(0.04)	(0.02)	(0.01)	(0.06)	(0.02)	(0.02)
Government higher experience	-2.64	1.51*	0.59	-1.77	-2.66***	0.48
	(1.85)	(0.83)	(0.50)	(2.20)	(0.87)	(0.74)
Female mayor	-6.72	0.28	1.09*	-9.35*	-2.70***	0.44
	(4.70)	(1.34)	(0.60)	(5.32)	(0.98)	(1.10)
Equality	-1.62	-0.82	-1.23*	3.09	0.71	0.36

	(2.27)	(1.17)	(0.73)	(2.44)	(0.65)	(0.87)
Small	13.00***	4.60*	-0.15	-40.52	8.17	2.08
	(4.55)	(2.38)	(1.52)	(31.70)	(6.04)	(2.28)
*Small*2015law*	-2.24	-6.66***	-1.94**	-8.47**	-0.01	-2.99**
	(2.76)	(1.25)	(0.81)	(3.58)	(1.04)	(1.20)
Mean forecast error	1.02	1.67	0.78	-6.31	-1.37	-1.69
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,720	1,720	1,513	1,720	1,720	1,720

Notes: Values are based on a standard Fixed Effects estimator from Equation (1). This is the same estimation reported in Table 1, Panel B. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure A1: Mean of the outcome variables over time for the whole sample



Notes: The dashed and solid lines mark the introduction of Order HAP/2105/2012 and HAP/2082/2014, respectively.

Table A4: Robustness tests

	Expenditures			Revenues		
	Total	Mandatory	Non-mandatory	Total	Own	External
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Alternative dependent variable						
*Small*2015law*	-9.03***	-5.37***	-1.87**	-8.98***	-1.68	-3.00**
	(2.86)	(1.34)	(0.85)	(3.18)	(1.64)	(1.33)
Mean forecast error	7.80	2.60	0.98	2.38	0.15	0.41
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,720	1,720	1,513	1,720	1,720	1,720
Panel B: Lagged dependent variable						
*Small*2015law*	-5.14***	-5.24***	-0.70	-10.90***	-0.86	-3.45***
	(1.89)	(1.17)	(0.48)	(3.55)	(0.71)	(0.97)
Mean forecast error	1.23	1.34	0.66	-7.23	-1.93	-2.09
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,548	1,548	1,361	1,548	1,548	1,548
Panel C: 2013-2019 sample						
*Small*2015law*	-8.25***	-4.84***	-0.59	-12.30***	-1.44	-3.79***
	(2.44)	(1.24)	(0.72)	(3.92)	(1.03)	(1.44)
Mean forecast error	0.72	0.85	0.11	-7.58	-2.78	-3.04
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,204	1,204	1,057	1,204	1,204	1,204
Panel D: Conditional on experience						
*Small*2015law*	-0.19	-4.49***	-1.42	-11.66***	0.13	-2.61
	(4.32)	(1.47)	(1.11)	(4.42)	(1.32)	(1.73)
*Small*2015law*highexp*	-4.80	-5.40*	-0.89	4.87	-0.60	-0.33
	(6.24)	(2.82)	(1.34)	(6.56)	(1.90)	(2.66)
Coefficient equality	[0.21]	[0.06]	[0.06]	[2.85]*	[0.07]	[0.31]
Mean forecast error	1.016	1.673	0.777	-6.314	-1.366	-1.692
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,720	1,720	1,513	1,720	1,720	1,720

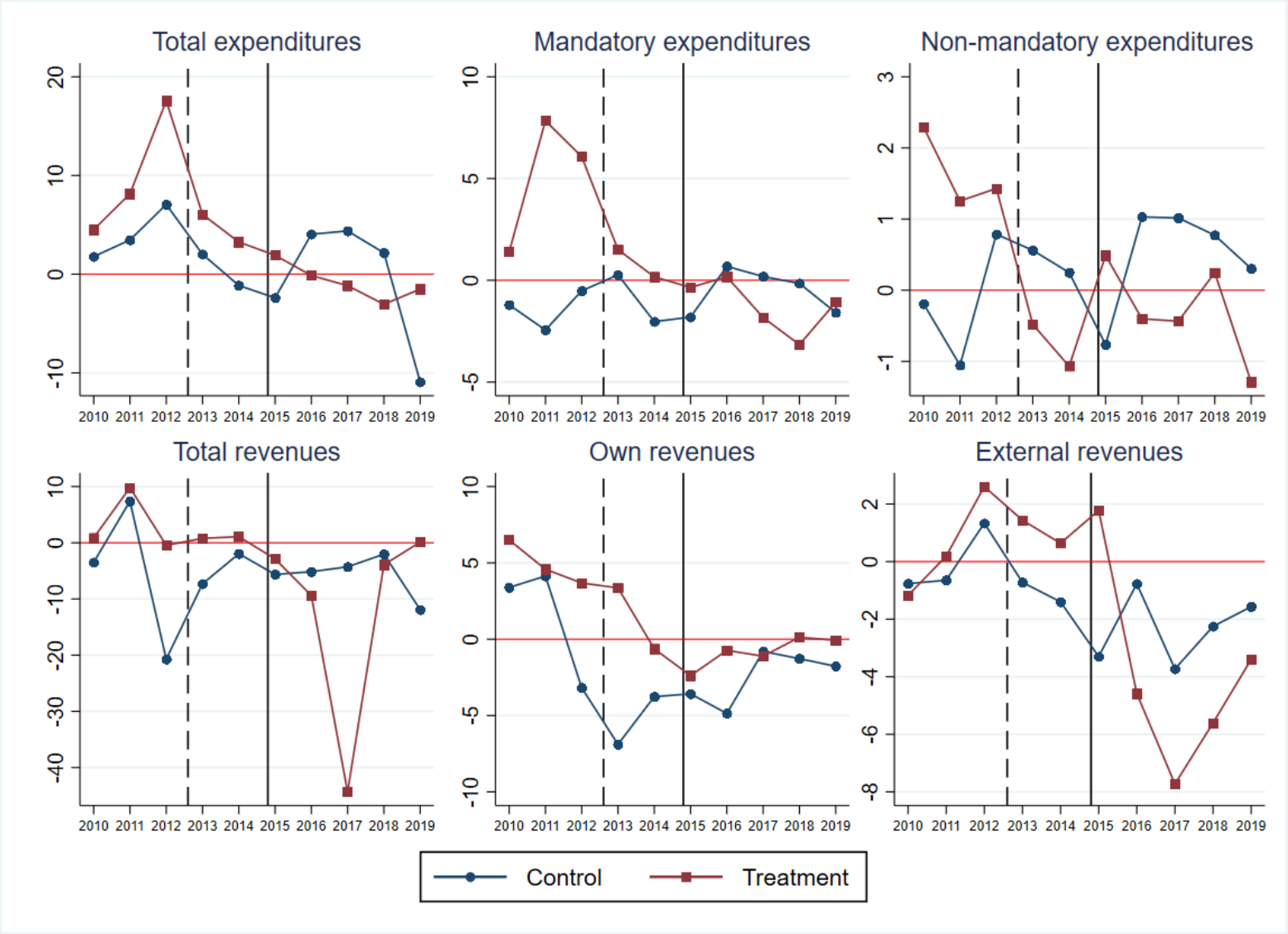
Notes: Values are based on a standard Fixed Effects estimator from Equation (1). Panel A uses the alternative dependent variable redefined as deviations over *actual* rather than *planned* budget. Panel B includes a lagged dependent variable in the regressors. Panel C restricts the sample to 2013 to 2019. Panel D interacts the treatment with *Government high experience*, and it also includes the F-test for equality of coefficients within brackets []. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Covariates before and after the 2015 transparency reform in the 'local' DID sample (2500 inhabitant bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Pre-			Post-		DiD
	Treatment	Control	Difference	Treatment	Control	Difference	(3)-(6)
Population	3422.80	6315.6	-2892.80***	3469.60	6345.63	-2876.03***	-16.77
	(75.72)	(78.72)	(110.31)	(74.84)	(75.84)	(107.64)	(154.18)
Unemployment rate	47.45	37.48	9.97***	34.83	28.36	6.47**	3.50
	(1.55)	(1.24)	(2.08)	(1.52)	(1.14)	(1.98)	(2.87)
Share of dependents	31.15	30.42	0.83**	32.45	31.60	0.85**	-0.13
	(0.23)	(0.20)	(0.32)	(0.20)	(0.20)	(0.28)	(0.43)
Share of immigration	16.04	15.93	0.11	13.09	13.08	0.01	0.10
	(0.61)	(0.40)	(0.77)	(0.49)	(0.36)	(0.64)	(1.01)
Rent	1283.37	1591.97	-308.60***	1552.38	1731.91	-179.53**	-129.07
	(40.65)	(51.59)	(64.78)	(45.89)	(45.53)	(65.58)	(92.16)
Left	0.17	0.16	0.01	0.24	0.24	0.00	0.01
	(0.04)	(0.04)	(0.05)	(0.04)	(0.05)	(0.06)	(0.08)
Parties in government	1.56	2.11	-0.55***	1.83	1.85	-0.02	-0.53**
	(0.10)	(0.14)	(0.16)	(0.11)	(0.11)	(0.16)	(0.23)
Turnout	75.66	72.42	3.24***	72.22	69.79	2.43***	0.81
	(0.34)	(0.40)	(0.52)	(0.36)	(0.43)	(0.56)	(0.76)
Mayor's higher education	0.22	0.52	-0.30***	0.25	0.45	-0.20**	-0.09
	(0.04)	(0.05)	(0.07)	(0.04)	(0.05)	(0.07)	(0.09)
Mayor's re-election	0.50	0.45	0.05	0.46	0.59	-0.12*	0.18*
	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.07)	(0.10)
Government higher education	19.81	27.43	-7.62**	17.94	25.90	-7.96**	0.34
	(2.18)	(2.52)	(3.32)	(2.01)	(2.53)	(3.19)	(4.61)
Government high experience	0.28	0.38	-0.10	0.38	0.61	-0.23**	-0.14
	(0.04)	(0.05)	(0.07)	(0.05)	(0.05)	(0.07)	(0.10)
Female mayor	0.17	0.39	-0.21***	0.17	0.25	-0.08	-0.13
	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)	(0.06)	(0.09)
Equality	0.48	0.45	0.03	0.57	0.63	-0.06	0.10
	(0.05)	(0.05)	(0.07)	(0.05)	(0.05)	(0.07)	(0.10)

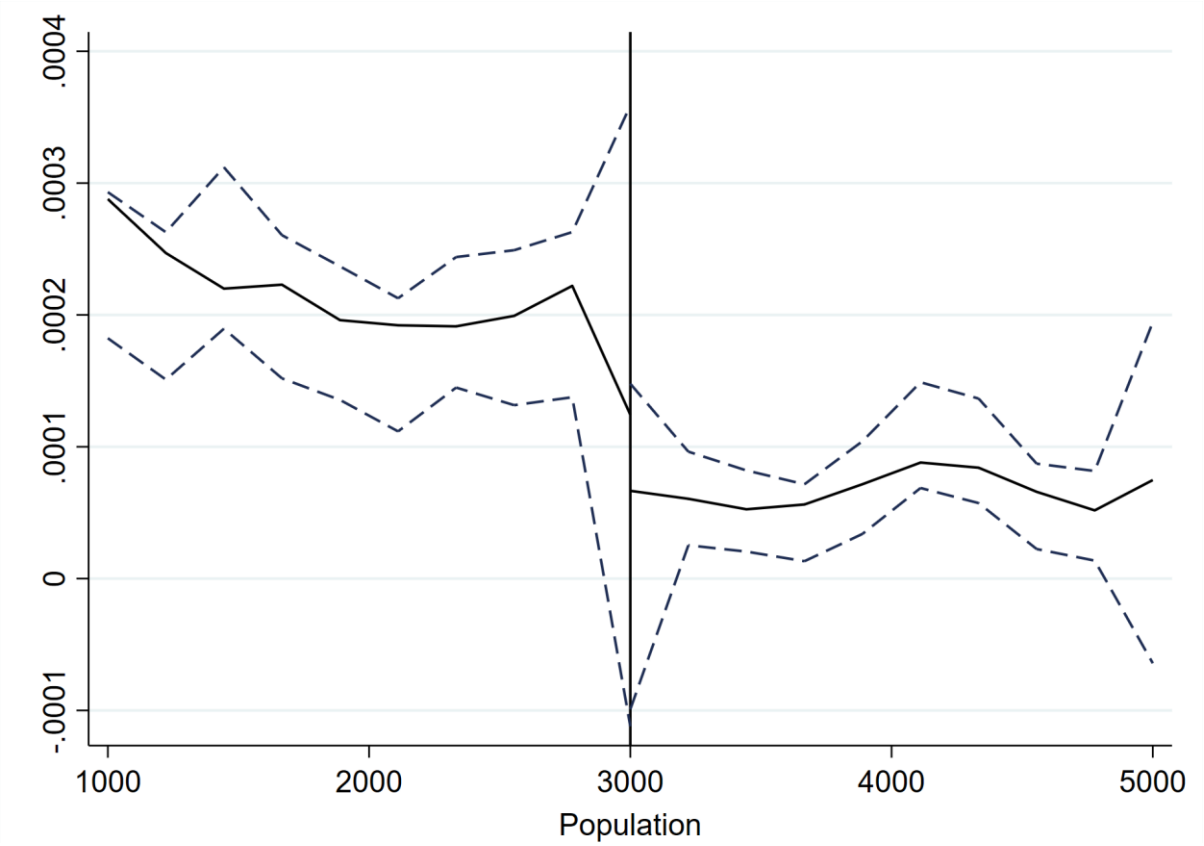
Notes: Difference indicates a t-test for equal means across the given treatment dimensions. *, **, and *** indicate significance at a 10%, 5%, and 1% level, respectively.

Figure A2: Mean of the outcome variables over time in the 'local' DID sample (2500 inhabitant bandwidth)



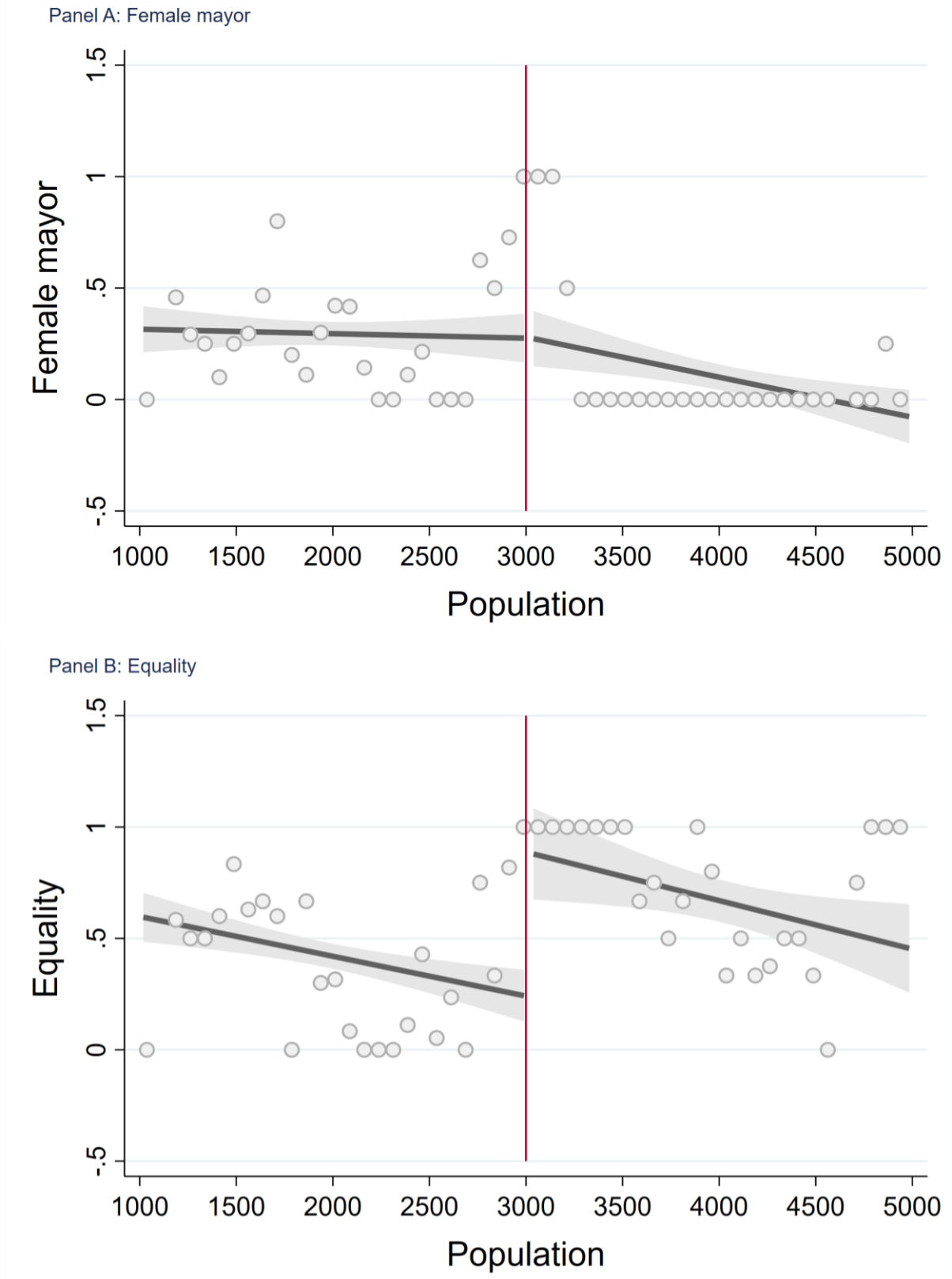
Notes: The dashed and solid lines mark the introduction of Order HAP/2105/2012 and HAP/2082/2014, respectively.

Figure A3: McCrary test – density around the 3,000 inhabitant threshold



Notes: McCrary (2008) test in densities for a second-order polynomial with evenly spaced bins. Dashed lines show 95% uniform confidence intervals.

Figure A4: The effect of the gender quota on having a female mayor and on equality



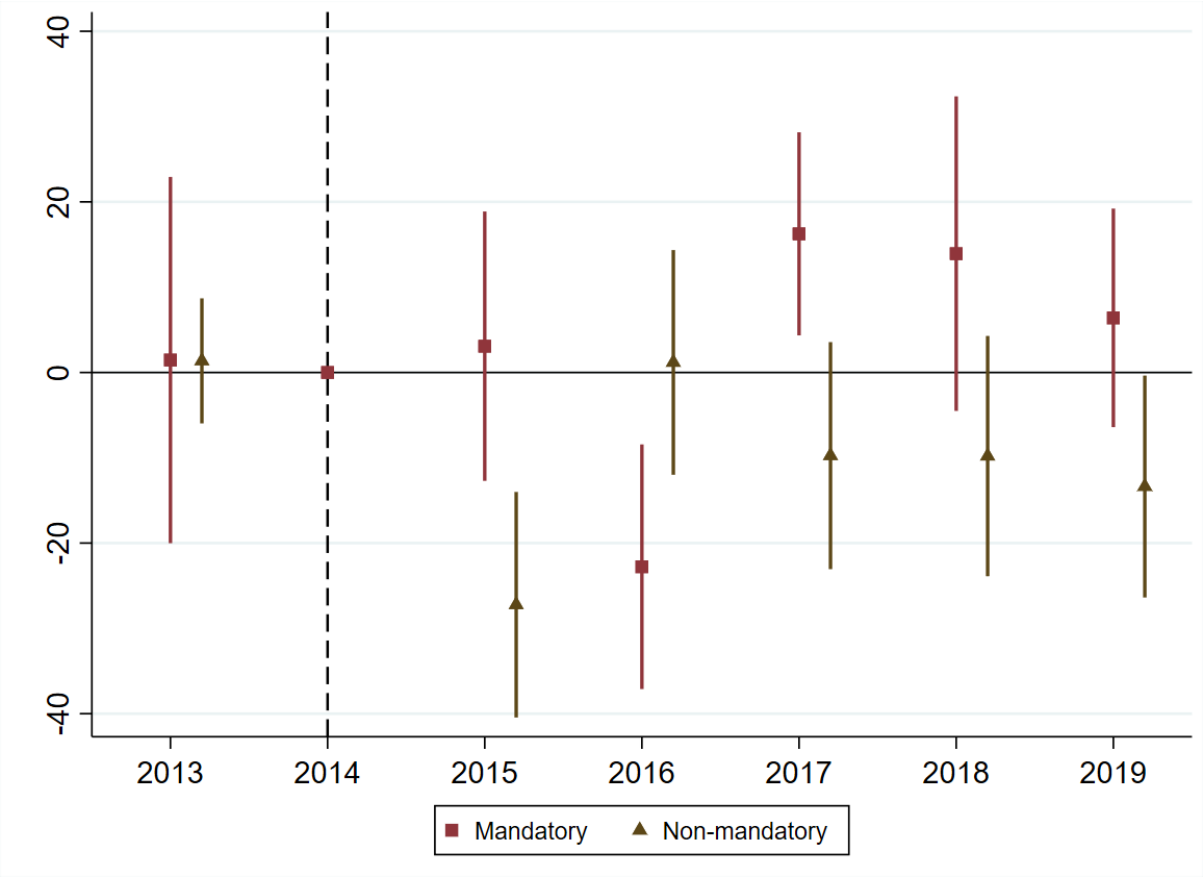
Notes: The top panel illustrates the probability of having a female mayor. The bottom panel plots the probability of having a gender-balanced government. Dots are averages in 75 inhabitant bins, and lines are regressions estimated on either side of the threshold separately. Shaded areas are the corresponding 95% confidence intervals.

Table A6: Gender differences in forecast errors ('Local' DiD)

	Expenditures			Revenues		
	Total	Mandatory	Non-mandatory	Total	Own	External
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 3,000 inhabitant threshold						
*Quota*2015law*	56.34	9.78**	-8.28*	64.51	12.15	-9.02
	(51.34)	(4.67)	(4.32)	(49.62)	(7.82)	(8.90)
Mean forecast error	-1.90	-0.07	-0.53	-6.92	-2.02	-3.02
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	500	500	500	500	500	500
Observations	82	82	82	82	82	82
Panel B: 1,900 inhabitant threshold						
*Placebo1900*2015law*	-23.89	3.23	-3.97	-10.69	13.80	3.60
	(31.60)	(9.49)	(5.27)	(35.54)	(10.68)	(8.34)
Mean forecast error	-1.88	0.93	-0.03	-10.75	-5.85	-5.68
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	500	500	500	500	500	500
Observations	123	123	123	123	123	123
Panel C: 4,100 inhabitant threshold						
*Placebo4100*2015law*	8.33	-4.23	-4.81	30.41	5.93	11.89
	(15.51)	(8.61)	(5.56)	(39.59)	(11.09)	(18.65)
Mean forecast error	4.73	-1.25	-0.30	-0.31	2.00	-1.19
Municipality & Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Bandwidth	500	500	500	500	500	500
Observations	53	53	53	53	53	53

Notes: Values are based on a 'Local' estimator from Equation (3) with both, municipality- and year-fixed effects. Panel A reports the results of relaxing transparency rules in municipalities below/above the 'true' *Quota* threshold (3,000). Panel B and C report the results of relaxing transparency rules in municipalities below/above artificial *Quota* thresholds (1,900 and 4,100). Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Figure A5: 'Local' event-study



Notes: I estimate the models reported in Table A6 (Panel A), but *2015law* is substituted by year dummies. The effect is set to 0 in the last year before the reform, and the resulting reference year 2014 is marked by a dashed vertical line. Each square and triangle in the graph shows the estimated effect, and the bars indicate 90% confidence intervals.