Joint Discussion Paper
Series in Economics

by the Universities of

Aachen · Gießen · Göttingen
Kassel · Marburg · Siegen

ISSN 1867-3678

No. 06-2023

Israel García and Bernd Hayo

Fiscal Reform in Spanish Municipalities: Gender Differences in Budgetary Adjustment

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

Fiscal Reform in Spanish Municipalities: Gender Differences in Budgetary Adjustment

Israel García and Bernd Hayo

(University of Marburg)

This version: 29 August 2023

Corresponding author:
Bernd Hayo
Marburg Centre for Institutional Economics (MACIE)
School of Business and Economics
University of Marburg
D-35032 Marburg
Germany
Phone: +49–6421–2823091
Email: hayo@wiwi.uni-marburg.de

Fiscal Reform in Spanish Municipalities: Gender Differences in Budgetary Adjustment

Abstract

Do gender differences matter for politicians' budgetary behaviour when confronted with an exogenous change in the institutional framework? After the 2013 Spanish municipal reform, municipalities with more than 20,000 inhabitants were no longer responsible for managing the provision of social services. Using a difference-in-differences estimator in a sample of municipalities from the Madrid region for 2010−2019, we compare gender differences in social services spending before and after the reform between municipalities below 20,000 inhabitants (control group) and above 20,000 inhabitants (treatment group). Although social spending was, on average, significantly reduced in the treatment group post-reform, we observe significant differences between municipalities conditional on the gender composition of local governments, i.e. council and mayor. Whereas male-dominated governments cut social expenditure by about 20% of the total budget, gender-balanced and female-dominated governments did not. Moreover, gender-balanced governments combined with female mayors increased social services spending by 40% more than gender-balanced governments combined with male mayors. This finding supports the claim that social spending is, on average, of particular importance to female politicians, as they are willing to bend the law to uphold their interests.

JEL Classification: C23, E61, D72, H75, I38, J16

Keywords: Gender, Difference–in–differences, Exogenous reform, Political budget cycles, Spanish municipalities, Madrid region

1. Introduction

Over the last decade, the introduction of explicit gender quotas helped increase the share of female politicians. Several empirical studies focus on the consequences of implementing gender quotas in politics. Their findings show that candidate list quotas (i) increase the pool of qualified female politicians (O'Brien and Rickne, 2016), (ii) improve the quality level of politicians (Baltrunaite et al., 2014; Besley et al., 2017), (iii) homogenise education levels of male and female politicians under closed-list proportional representation (Profeta and Woodhouse, 2022), (iv) increase electoral turnout, and (v) help lessen negative stereotypes against women (De Paola et al., 2010; 2014). The latter result is consistent with Esteve-Volart and Bagues (2012) and Casas-Arce and Saiz (2015), who report that prior quota, Spanish party leaders were not maximising electoral success, as they included fewer women in the candidate lists than voters would have preferred.

A widespread assumption in the literature is that women hold distinctive social preferences and favour specific fiscal budget items, mainly social, health, and education spending. Bagues and Campa (2021), among others, call these types of budget items 'female' expenditures. Empirical evidence supporting this conjecture is mixed, as few researchers find that female political representation impacts policy choices (Hessami and da Fonseca, 2020). Studies exploiting exogenous variation in women's representation in politics to estimate the impact of gender report that an increase in the share of female politicians has no impact on the composition of public spending (Ferreira and Gyourko, 2014; Bagues and Campa, 2021). Geys and Sørensen (2019, 3) note that in most high-income countries, there are substantial institutional and budgetary constraints, especially at the local level, which suggests that higher female representation may not easily generate notable swings in public policies.

We take a new perspective on the influence of gender quotas on political outcomes by examining whether gender differences matter for politicians' behaviour when confronted with an exogenous change in the existing institutional framework. We exploit the introduction of a new law in Spain aimed at reducing the scope of social services at the municipal level. Our general hypothesis is that exogenously changing the spending framework of a budget item considered to be of particular interest to women will, conditional on politicians' gender, cause a different budgetary impact in the affected municipalities. Specifically, we conjecture that the combination of a female mayor and a gender-balanced government causes the social services budget item to react (1) very little, if at all, in absolute terms and (2) certainly less than the adjustment taking place in more male-dominated districts.

In December 2013, the 'Local Government Rationalisation and Sustainability Act' (No. 27/2013) was introduced in Spain. However, as the initial budget was planned in advance, the 2015 budget was the first affected by the reform. The Act intended to clarify municipal competences and eliminate any overlap in responsibility between local and regional or central governments. In Article 26, the Act sets out new mandatory competences for local governments, including a significant shift in responsibility for social services from municipal to regional governments. Before the law's introduction, municipalities with more than 20,000 inhabitants were responsible for managing social services. After the reform, the municipal responsibility is reduced to identifying situations where social assistance is needed. Hence, while still a mandatory service, it is now highly limited. The reform-induced reduction of municipal competences suggests that local governments will reallocate fiscal spending away from social expenditure. Moreover, due to the particular interest female politicians appear to have in social spending, we expect this reform to trigger asymmetric responses across municipalities in the form of relatively more substantial cuts in male-dominated local governments.

To study the budgetary effects of Act 27/2013, we assemble data from the Madrid region (Comunidad Autónoma de Madrid) from 2010−2019. Municipalities in the Madrid region are subject to a homogenous set of budget rules. Furthermore, in July 2014, the Madrid region introduced a new law (No. 1/2014), forcing local governments to adjust to the Act. This law aims to avoid overlapping competences between jurisdictions while maintaining a public service guarantee. More generally, 2010−2019 is characterised by balanced local budgets and constraints on total expenditures due to the Organic Law 2012. This sample period provides the same number of years before and after the reform and covers three different electoral terms (elections took place in 2011, 2015, and 2019).

We run municipal-level regressions using a difference-in-differences (DiD) estimator to compare gender differences in social services spending before and after the reform. The control group are those municipalities below 20,000 inhabitants, the mandatory tasks of which do not include social services, whereas the treatment group are those municipalities with 20,000 or more inhabitants where social services are part of the mandatory tasks. Although the reform does not eliminate social spending from being a mandatory competence, its scope is reduced substantially.

First, we compare social services spending before and after the reform. Our DiD estimates show that the reform reduced social spending by more than 20%, which suggests that the reform had the intended effect across all municipalities *on average*.

Second, we condition the analysis on politicians' sex. In the whole sample and as long as the government is male-dominated, treated municipalities spend significantly less on social services than the control group. The effect ranges from a nearly 50% reduction of average social spending for governments with a less than 30% share of female politicians to a 20% reduction for governments with a more than 40% share of female politicians. However, we find no statistically significant difference between municipalities led by a male or female mayor. These results suggest that the differences between untreated and treated municipalities are driven by those governed by male-dominated councils.

Third, the sample is split between male-dominated and gender-balanced governments. In the male-dominated government sample, treated municipalities spend less than untreated municipalities after the introduction of the reform. When the effect is conditioned on the mayor's sex, we discover a significant increase in social spending, averaging 40% for municipalities characterised by gender-balanced governments and female mayors compared to those characterised by gender-balanced governments and male mayors.

According to García and Hayo (2022), budget decisions are linked to the electoral term in the local governments of the Madrid region. When considering the impact of elections, we still find that, on average, treated municipalities significantly reduced their share of social spending in the aftermath of the 2015 reform. However, the effect is reversed in pre-electoral years, which implies that treated municipalities tend to increase their share of social spending before elections.

Overall, our analysis supports our initial hypothesis that a budgetary reform can trigger different responses conditional on politicians' sex. Although social spending was, on average, significantly reduced in the treated municipalities after the reform, the effect is mainly driven by male-dominated governments. Thus, despite the change in the fiscal framework, gender-balanced governments appear to be relatively more reluctant to reduce social spending. Mayors' sex is only relevant in a subsample of gender-balanced governments. In addition, we show that under specific circumstances, such as having a preference for social spending, female politicians may behave as opportunistically motivated as male politicians. Hence, once institutional and budgetary constraints are accounted for, higher female representation can generate notable election-related swings in public policies.

The remainder of the paper is organised as follows. Section 2 discusses the related literature, Section 3 describes the institutional context, and Section 4 sets out the identification strategy and our dataset. The empirical strategy and the estimation results are presented in Section 5, and Section 6 concludes.

2. Related literature

There is substantial evidence that women have different social preferences than men. Generally, redistribution and equality seem stronger for women (Corneo and Grüner, 2002; Alesina and La Ferrara, 2005). Moreover, using a controlled experiment comparing matrilineal and patriarchal communities, Andersen et al. (2008) and Gneezy et al. (2009) report that competitive behaviour and public good provision are affected by sex. Oswald and Powdthavee (2010) conjecture that having daughters makes people more likely to vote for left-wing parties. Women's enfranchisement and its impact on the size and composition of government have been investigated, too. Lott and Kenny (1999) and Bertocchi (2011) find that the size of the government increased when the franchise was expanded to women, typically through higher expenditures on health, education, and social issues (Aidt et al., 2006; Aidt and Dallal, 2008). Funk and Gathmann (2015) use data from a survey in Switzerland to report that in specific areas such as health, defence, environmental issues, and welfare spending, men and women show distinct predilections. Employing representative survey data for Germany, Hayo and Neumeier (2019) provide evidence that women are less enthusiastic about fiscal consolidation than men.

A related line of research studies the discrepancy in preferences between male and female politicians. Washington (2008) shows that parenting daughters increases a congressperson's propensity to vote liberally on reproductive rights issues and support women's issues. Slegten and Heyndels (2019) provide survey evidence for a more left-wing position of female politicians. Baskaran and Hessami (2019), Weeks (2019), and Lippmann (2022) find that female political representation increases political discussions and amendments on issues such as childcare support and social justice.

Several studies focus on the implications of implementing women quotas in politics (Hessami and da Fonseca, 2020). Empirical findings demonstrate that candidate list quotas not only increase the pool of qualified female politicians (O'Brien and Rickne, 2016) but improve the overall quality level of male and female politicians by driving out lower-competent male politicians (Baltrunaite et al., 2014; Besley et al., 2017). Additionally, these quotas homogenise the education levels of male and female politicians under closed-list proportional representation (Profeta and Woodhouse, 2022). De Paola et al. (2010, 2014) show how quotas increase electoral participation and help reduce negative stereotypes against women. This result is consistent with the findings of Esteve-Volart and Bagues (2012) and Casas-Arce and Saiz (2015), which show that prior quota, party leaders in Spain were not maximising electoral success, as they included fewer women in the candidate lists than voters would have preferred.

Reflecting the fact that the share of female politicians has increased in recent years, several empirical studies focus on the influence of gender differences in policymaking. Using a quasi-experimental setup in India, Chattopadhyay and Duflo (2004), Clots-Figueras (2011, 2012), and Bhalotra and Clots-Figueras (2014) provide evidence of gender differences in fiscal expenditures on health and education, which are considered to be of particular interest to women.

Studies on gender differences in fiscal spending at local governmental levels in more prosperous economies yield mixed results. On the one hand, Svaleryd (2009), Funk and Gathmann (2015), Braga and Scervini (2017), Clayton and Zetterberg (2018), Funk and Philips (2019), and Casarico et al. (2022)

find gender differences concerning spending on childcare, health, education, and social assistance. On the other hand, Gagliarducci and Paserman (2012), Ferreira and Gyourko (2014), Geys and Sørensen (2019), and Bagues and Campa (2021) report that the composition of public expenditure is unaffected by different shares of female politicians.

Using municipal data on Spain, Navarro-Galera et al. (2017), Cabaleiro-Casal and Buch-Gómez (2018), Hernández-Nicolás et al. (2018), Cabaleiro-Casal and Buch-Gómez (2020, 2021), and Balaguer-Coll and Ivanova-Toneva (2021) investigate the effect of female mayors and councillors on debt, financial stability, and ‘fiscal austerity’. Generally, an increase in the share of female councillors is correlated with lower debt and a better financial situation of the municipality, but results are mixed concerning total expenditure and the impact of female mayors.

Gender differences in political leadership also appear to occur regarding electorally-motivated spending. Accettura and Profeta (2022) and García and Hayo (2022) study gender differences in the context of political budget cycles. Whereas Accettura and Profeta (2022) report (primarily weak) evidence that male mayors are more likely to engage in pre-electoral spending, García and Hayo (2022) show that female mayors use pre-electoral spending strategically, although to a lesser extent than their male counterparts do. Moreover, they emphasise that the share of women in government plays an essential moderating role in election-oriented spending patterns.

3. Institutional context

3.1 Electoral system

Local elections are held throughout the country every four years on the same day. Councillors are elected through a proportional representation system based on closed lists.[1] There are as many electoral ballots as there are parties in each municipality, and each ballot includes as many candidates as the number of possible councillors. To ensure proper representation, the number of elected councillors is computed according to the d’Hondt law, combined with a 5% threshold to avoid a 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, and, in principle, council members serve four-year terms (Organic Law 5/1985, ‘General Electoral Regime’). The mayor proposes initiatives and regulations, which are passed, or not, by majority voting in the council. The mayor has control over the municipality’s executive functions and presents and

[1] Municipalities with 250 (or less) inhabitants use an open-list system.

explains the municipal budget proposal to the council.[2] The council is responsible for monitoring the municipality's activities and approving the budget and its possible amendments.

In March 2007, the Equality Act mandated gender-balanced candidate lists. According to the Act, at least 40% of the candidates on an electoral list must be female, and at least 40% must be male. This quota applies both to the entire party list and to 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 at least six men, plus at least two men and two women within the first five positions and within positions 6 to 10. Parties whose candidate lists do not fulfil these requirements are not allowed to participate in the elections. The quota was introduced in 2007 in municipalities with more than 5,000 inhabitants; in 2011, it was extended to municipalities with more than 3,000 inhabitants.

3.2 Local public finance in Spain

In Spain, local governments are responsible for a significant number of 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. Dependent on population size, municipalities must provide different levels of basic services:[3]

- all municipalities: public lighting, cemeteries, waste collection, public cleaning, drinking water supply, sewer system, access to urban areas, road surfacing, and food and drink control
- municipalities with more than 5,000 inhabitants: all of the above, plus public parks, public libraries, and market and waste management
- municipalities with more than 20,000 inhabitants: all of the above, plus civil defence, social services, fire safety, sports facilities, and slaughterhouse
- municipalities with more than 50,000 inhabitants: all of the above, plus public transport and environmental protection

In 2001, Spain passed the Law of Budgetary Stability to accommodate the European Monetary Union rules on public finances. Consequently, local governments must comply with the Balanced Budget Rule (BBR), which states 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. Furthermore, local governments may incur deficits only under special circumstances and with the authorisation of the Ministry of Finance. From 2009 onwards, local governments suffered a severe

[2] Sweeting (2009) provides a detailed discussion of the mayor's role in Spain.

[3] According to Law 27/2013 (27 December 2013), food and drink control, markets, and slaughterhouses are not part of mandatory minimum services from the 2015 budget onwards.

worsening of public finances, and, to maintain budgetary stability, the Spanish government adopted austerity programmes that included tax increases and public spending reductions. In 2011, public budget stability was anchored in the Spanish Constitution (Article 135). Under this provision, local governments were required to adhere to the BBR and debt repayment took priority over any other expenditure. A year later, the Spanish parliament approved Organic Law 2/2012, 'Budgetary Stability and Financial Sustainability', to operationalise the budget stability obligation implemented in the Constitution. This law further tightens municipalities' fiscal limits by regulating government expenditures.

3.3 Local budget and the Act 27/2013

In December 2013, the 'Local Government Rationalisation and Sustainability Act' (No. 27/2013) was introduced to ensure that local governments comply with the rules and regulations previously set out in the Spanish Constitution (Article 135) and Organic Law 2/2012. The intention of Act 27/2013 was to clarify municipal competences and eliminate any overlap in the responsibilities of local and regional or central governments. In Article 25, paragraphs 3, 4, and 5, the law establishes the basic rules for municipalities to fulfil their competences. First, an economic criterion is imposed: any function assigned to the municipality must be accompanied by the provision of sufficient resources to fulfil that function so that the municipality's budgetary stability is not compromised. Second, a decentralisation of competences should only be undertaken as long as it does not undermine the efficient allocation of public resources. Third, doubling competences is strictly forbidden so that no specific function can be assigned simultaneously to two public administrations.

In Article 26, the Act establishes the new scope of mandatory competences for local governments, including a notable reduction in municipal social service responsibilities in favour of the regional government. Before the law's introduction, municipalities with more than 20,000 inhabitants were responsible for managing the social services provision. After the reform, this social services function is no longer strictly speaking a service but a simple identification of situations where social assistance is needed. While still a mandatory function, it is now highly limited, and any expenditure exceeding the amount necessary to achieve the mandatory minimum level is subject to BBR and the other fiscal rules presented in the previous paragraph.

4. Identification and data

In order to analyse gender differences in policymaking, we exploit the introduction of Act 27/2013 and focus on social services and its extensive modification by Article 26. Using a DiD estimator, we conduct municipal-level regressions to compare gender differences in social services spending before and after the reform. The control group are municipalities below 20,000 inhabitants, which do not have to provide social services, whereas the treatment group are municipalities with 20,000 or more inhabitants for which social services are mandatory. As we explained in Section 3.2, at this threshold, both mandatory expenditures and the amount of transfers received by the local government change. However, there have been no further modifications to these rules since the reform. As shown in Section 2, in high-income countries, empirical evidence on gender differences and the influence of politicians' sex on policy choices is mixed. This could be due to substantial budgetary and institutional constraints, particularly at the local level, which create barriers to notable swings in public policy (Geys and Sørensen, 2019).

Our analysis does not rely on an exogenous change in women's representation in politics, the standard procedure. However, we examine gender differences in politicians' behaviour when confronted with an exogenous change in the existing institutional framework. Although Act 27/2013 keeps social spending as a mandatory task, its scope is clearly reduced.

Our general hypothesis reflects current literature and postulates that exogenously changing the spending framework of a budget item considered to be of particular interest to women will, conditional on politicians' gender, cause a different budgetary impact in the affected municipalities. Specifically, we conjecture that the combination of a female mayor and a gender-balanced government causes the social services budget item to react (1) very little, if at all, in absolute terms and (2) certainly less than the adjustment taking place in more male-dominated districts.

The sample comprises 175 municipalities from the Madrid region (Comunidad Autónoma de Madrid) and covers 2010−2019. For several reasons, the Madrid region can be considered an interesting case. First, available information at the local level is generally not homogeneous between regions. Thus, building a comparable database using data from municipalities located in different regions is challenging. Second, Madrid has no supra-municipal authority between municipalities and regional government, which ensures homogeneity in terms of legal requirements concerning public services provision and grants.[4] Third, in light of Act 27/2013, the Madrid region introduced Law 1/2014 in July 2014, intending to avoid the doubling of competences while guaranteeing the provision of all mandatory public services. At the same time, the Madrid region is very diverse; it includes a *very large*

[4] 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 must provide some mandatory services according to their population size (see Section 3). The non-mandatory services are provided by either the regional or the central government. Article 36 of the local administrative law states that the provincial administration is in charge of coordinating and establishing mandatory municipal services. According to the territorial administration, it is possible that two similarly sized municipalities, which belong to the same region but different provinces, could have different standards in their mandatory services. Madrid is a region with only one province; consequently, there cannot be any variability in mandatory services across the various municipalities.

municipality, some *large* ones, and a considerable number of towns and villages. Most municipalities in our sample (55%) have less than 5,000 inhabitants, whereas 13% have more than 50,000. Note that, compared to the rest of the country, our sample region under-represents small municipalities and over-represents large ones.[5] This over-representation of large municipalities provides us with a greater number of treated 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 extensively modified, thus making a comparison of most expenditure items before and after 2010 practically impossible. 2010−2019 is characterised by balanced local budgets and constraints on total expenditures due to Organic Law 2012 and Act 27/2013. This period provides an equal number of years before and after the reform and covers three different electoral terms from elections that took place in 2011, 2015, and 2019.

Following García and Hayo (2022), we used planned budgets rather than actual budget data. Since planned budgets are published before the actual budget period, they can be understood as a forward-looking signal that reveals the preferences and/or competence of the incumbent local government. The council must approve the initial budget, and its potential successive modifications, by majority. Thus, the planned budget works well as a signalling device, whereas the actual budget is a better proxy for the concrete provision of public services. Since we base our analysis on signalling theory, we believe the planned budget to be the appropriate measure.

We extracted this data from the Ministry of Finance's CONPREL database and merged it with data collected from municipalities' archives.[6] All budget items are measured as a share of the total budget in per cent. From the expenditure side, we obtain our variable of interest: *Social services spending*. We collect three variables from the revenue side: *Own revenues*, *Current transfers,* and *Capital transfers*. *Own revenues* capture each municipality's degree of fiscal autonomy. *Current transfers* are unconditional transfers from the region based on population size, whereas *Capital transfers* can only be used to finance investment projects and are conditional on fulfilling additional requirements. There is evidence that *Current transfers* are biased towards municipalities with large populations since these municipalities receive more than their population share would suggest, whereas smaller

[5] According to the Spanish Statistical Institute, in 2019, out of 8131 municipalities, 83% had less than 5,000 inhabitants and 5% more than 50,000 inhabitants.
[6] https://serviciostelematicosext.hacienda.gob.es/SGFAL/CONPREL.

municipalities profit over-proportionately from *Capital transfers* (Solé-Ollé and Bosch, 2005; Solé-Ollé and Sorribas-Navarro, 2008, 2012; Curto-Grau et al., 2018).

Socioeconomic variables were collected from the Statistical Institute of the Community of Madrid.[7] According to the social services reference catalogue: 'social services should be aimed at population groups that, due to their singular characteristics or their vulnerability, need special attention'. These population groups comprise people with disabilities, the elderly, young people at risk of social exclusion, homeless people, and immigrants (Cerreda, 2015). To account for these groups in our analysis, we include various indicators, all of which are measured in per cent of the population. *Rent* is a measure of municipal income.[8] *Debt*, measured as outstanding debt at the end of the fiscal year, controls for the municipality's level of indebtedness, which is relevant in the Spanish budget process (Cabaleiro-Casal and Buch-Gómez, 2018; Cabaleiro-Casal and Buch-Gómez, 2021; Balaguer-Coll and Ivanova-Toneva, 2021). *Debt* and *Rent* are expressed in per cent of total spending.

Political data were retrieved from the Ministry of the Interior's Database of Electoral Results.[9] To create a consistent dataset referring to the governing body rather than the whole council, we combine these variables with specific information on each municipality. We define the governing body as those councillors who are in the governing majority, either via one party's absolute majority or in a coalition (see García and Hayo, 2022).[10] *Left* and *Right* measure the government's ideological orientation. When the mayor belongs to the Popular Party or Citizens, or one of the two parties is the prominent member in a coalition, the municipality is defined as right-wing oriented. Socialist Party or Left United indicate left-oriented municipalities. The rest of the municipalities are administered by local parties, which are hard to place on a left-right scale. *Parties in government* capture the degree of government fragmentation. *Mayor's age* is measured in years. *Mayor's primary*, *secondary*, and *higher education* are three dummies that take the value 1 when the mayor has obtained primary, secondary, and university education, respectively. Our gender variables of interest are *Female mayor*, a dummy variable taking the value 1 in case of a female mayor and 0 otherwise, and *Female government*, measured in per cent, which is computed using the share of women in the governing body as defined above.

[7] http://gestiona.madrid.org/desvan/Inicio.icm?enlace=almudena

[8] This variable is constructed based on information provided by tax authorities (as the main input), plus Information on earnings, wealth (capital and non-capital), rents, social payments, and transfers in each municipality. The weights of each of these components are adjusted according to different factors, such as the number of households, number of declarants, age of the population, percentage of rents from non-working earnings, and a socioeconomic indicator for each municipality.

[9] http://www.infoelectoral.mir.es/infoelectoral/min/

[10] When a party does not have an absolute majority and a coalition is not reached, the whole council is considered the governing body.

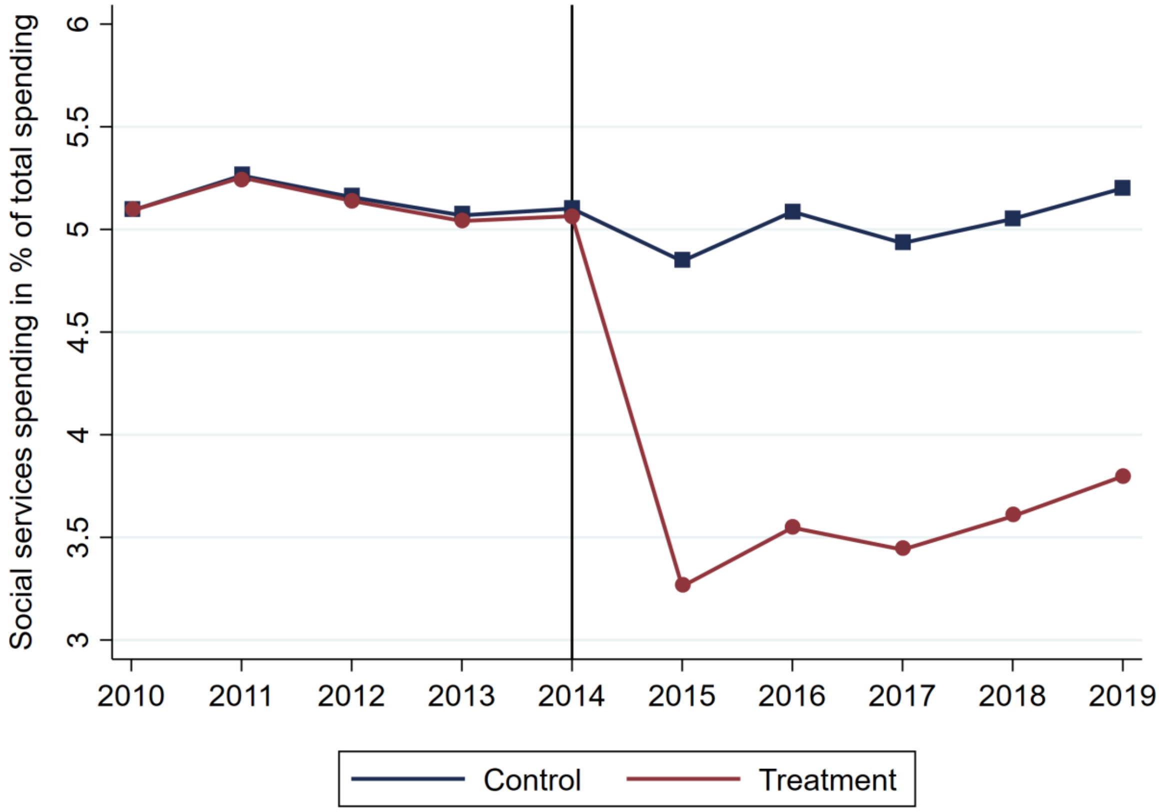
Our treatment variables are two dummies and their interaction, the DiD estimate. *2015law* is a dummy that takes the value 1 for each year from 2015 onwards and 0 otherwise. Act 27/2013 was approved at the end of December 2013, and it was too late to implement the changes in the planned budget for 2014. Thus, the first planned budget affected is the one referring to 2015. *Treated* is a dummy variable that takes the value 1 for municipalities with 20,000 or more inhabitants and 0 otherwise. To account for the electoral term, we include three dummies: *Pre-election, Election, and Post-election.*

Before turning to our empirical strategy, we provide evidence that a DiD approach is suitable in our case. Administratively, the Madrid region is divided into 179 municipalities. We restrict our sample to the 175 municipalities without missing observations, which yields a panel dataset containing 1,750 observations. In our sample, 80% of municipalities (140) belong to the control group, whereas 20% of municipalities (35) are part of the treatment group. A possible concern is the self-sorting of municipalities into control or treatment groups according to their social services spending preferences. In our case, the assignment into the groups 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, four municipalities fell into the interval of 19,000 to 21,000 inhabitants; amongst these, one municipality 'jumped' above the threshold of 20,000 inhabitants. This result is in line with Foremny et al. (2017) and Bagues and Campa (2021), who show that Spanish municipalities exhibited some sorting around the threshold of 5,000 inhabitants before 2005, but they found no evidence of sorting after 2005. They conclude that sorting at higher population thresholds, e.g. 20,000 inhabitants, is much more difficult since the Spanish Statistical Office seems to monitor larger municipalities more intensely.

Figure 1 depicts the development of average social services spending in per cent for the control and treatment groups. These are the annual observed means, conditional on time fixed effects and group time trends and control and treatment group constrained to be equal in the first period. *Social services spending* in municipalities below 20,000 inhabitants (control group) shows a similar pattern before and after the reform. The treatment group experienced a downward adjustment after Act 27/2013 came into force. Hence, the primary adjustment in planned social services expenditures happened right after the reform came into force in 2015. Afterwards, the control and treatment groups show a similar trend.

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. Table A2 of the Appendix reports averages of our covariates in the two groups before and after the budgetary reform as well as their differences. In the pre-treatment period, columns 1, 2, and 3, both groups differ in almost all characteristics. This is not surprising, however, since the classification as treated versus not treated is based on population size, and one would expect that differently-sized municipalities are dissimilar.

Figure 1: Linear-trends model of social services expenditure in relation to total expenditure (in %)



More important for our purposes is the fact that these differences are stable during the whole period analysed (see columns 4, 5, and 6 of Table A2 for the post-treatment period), and thus, differences in social services spending can be solely attributed to the reform and not to changes within control or treatment group. At a 5% level of significance, we discover statistical differences in *Rent, Old unemployment rate*, *Parties in government, and Share of women in government*. Given our research question, the difference in the latter variable may be a reason for concern. As stated in Section 3, in March 2007, the Equality Act was enacted in municipalities with more than 5,000 inhabitants to increase gender-balanced candidate lists. In 2011, it was extended to municipalities with more than 3,000 inhabitants. If we account for this change in the control group, differences in the share of women in government before and after the reform are no longer significant.

5. Empirical strategy and results

5.1 Method

To show that our quasi-experimental design is the driving force of results, we commence our analysis by estimating a classic DiD, which includes neither controls nor a lagged dependent variable. However, we do not think that this is necessarily the best specification. Although the correlation between the controls and the treatment should be zero by design, Leamer (2010) raises doubts that the experimental randomisation of the treatment eliminates the requirement to include additional controls in the equation. In the case of finite samples, the 'zero-by-design' argument may not hold, as the correlations between the randomised treatment and the controls, by chance, tend to be non-zero. Moreover, modelling the data generation process more precisely through the inclusion of control variables reduces the error of estimation, which means the standard errors of the treatment effects decrease, too, and we can estimate the coefficient of interest more efficiently (Hayo 2018). Finally, dynamic models allow differentiating between the short- and long-term impacts of explanatory variables, which is of particular interest when studying public budget cycles.

Since our dependent variable, *Social services spending,* is characterised by a considerable degree of persistence, we consider a lagged dependent variable. In addition, we include a set of control variables and the DiD estimate of interest in our preferred specification. The baseline DiD specification is:

$$y_{i,t} = \delta y_{i,t-1} + \sum_{j=0}^{1} \varphi' X_{i,t-j} + \gamma Treated_{i,t} + \beta(Treated_i * 2015law_t) + \mu_i + \tau_t + \varepsilon_{i,t} \qquad (1)$$

where i is the municipality index, t is the year index, $y_{i,t}$ is *Social services spending*, $X_{i,t-j}$ 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. The j index represents the fact that we allow for some dynamics, not only with respect to the lagged dependent variable but also regarding the control variables. *Treated* allows us to capture the unobserved time-invariant characteristics that may differ across municipalities in the two groups. *Treated*2015law* gives the DiD estimate for the reform's effect on social services expenditure. To avoid perfect collinearity with the time fixed effects, we do not include *2015law*.

Given that our data cover some years before and after the reform, we can also examine the reform impact over time using event-study plots. We substitute variable *2015law* by year dummies and estimate our reported models again. We use 2014, the last year before the reform, as our reference year, and we compare annual effects before and after the reform.

For a small number of time-series observations *T*, considering a lagged dependent variable causes the panel data fixed effect (FE) estimator to become inconsistent (Nickell 1981; Kiviet 1995). In our analysis, T equals 10, and hence, the resulting bias could be non-negligible. In light of this, we opt for using a bias-corrected FE estimator. Recently, Breitung et al. (2022) proposed a bias-corrected estimator that is easier to implement than likelihood-based estimators and has an advantage over more 'classical' bias-corrected methods in that heteroscedasticity and cross-sectional dependence are accounted for when computing cluster-robust standard errors. Small-sample Monte Carlo simulations of the Breitung et al. estimator suggest that it tends to outperform 'classical' bias-corrected FE and the general method of moments estimators in terms of bias.[11]

Using the Breitung et al. estimator allows for including many control variables and accounting for a reasonably dynamic influence of the explanatory variables. We recognise the possibility of inertia in the adjustment of expenditures to the budget's revenue side by including current and lagged values of *Own revenues*, *Current transfers,* and *Capital transfers.* All socio-economic variables are used only in lagged form since current values are published well after the budget has been approved. Political variables are also included in the lagged form, as the planned budget in year t was proposed by the governing body in $t-1$.

The disadvantage of using so many control variables is a loss in estimation efficiency, and our general specification is likely over-parameterised. To address this issue, we resort to general-to-specific modelling (Hendry, 1993). Applying a consistent testing-down procedure at a nominal significance level of 10% allows us to arrive at a more efficiently estimated model (reduced model based on a smaller sample) while accounting for collinearity and standard-error-reducing complementarity (Hayo, 2018).[12] Several variables with missing values do not survive the testing-down procedure. When estimating the reduced models, we can use them to increase the sample size, which allows us to conduct a (partial) out-of-sample test of the models' stability. In the main text, we focus our discussion on these reduced models estimated with the maximum number of available observations (reduced model based on larger sample), as they are, on the one hand, admissible statistical representations of the general models and, on the other hand, more efficiently estimated than the other two types of models.[13] Moreover, given that the number of municipalities remains unchanged, comparisons with results obtained in the classic DiD estimation are straightforward.

In order to condition the analysis on politicians' sex, the DiD term from the baseline model in equation (1) is separately interacted with either a female mayor dummy (*Female mayor*) or the share of female

[11] For a deeper discussion of the topic, see Breitung et al. (2022) and García and Hayo (2022).

[12] In the general models, the DiD estimates of interest tend to show lower p-values than the ones obtained by the simplified models. However, standard errors are larger in the former than in the latter. Increasing estimation efficiency, rather than minimising p-values, is the reason for deriving the reduced models.

[13] All estimates for the general and the reduced model, with the same set of observations used for estimating the general model, are reported in Appendix B.

politicians in the government (*Female government*). Afterwards, we study the effect of both gender dimensions jointly. Reflecting female politicians' alleged preference for social spending, we specifically conjecture that the combination of a female mayor and a gender-balanced government causes the social services budget item to react (1) very little, if at all, in absolute terms and (2) certainly less than the adjustment taking place in more male-dominated districts.

In the Spanish local electoral system, candidates who run for mayor are the most visible and, very frequently, the only candidates that voters even recognise (Sweeting, 2009; Bagues and Campa, 2021). Therefore, a possible threat to our identification strategy is that deciding to vote for a male or female mayor in a given municipality might be correlated with the voter’s desired level of social services. To address this issue, we perform the empirical analysis in a restricted sample of municipalities where a female (or male) mayor was elected in both the 2011 and 2015 elections; that is, the governing body was headed consistently by one sex. We would argue that, at least in these municipalities, the reform's effect can be interpreted as causal and not just as a by-product of a possible change in social spending preferences after the 2015 election. Implementing this restriction results in a panel dataset comprising 1,370 observations, with 80% (110) of municipalities in the control group and 20% (27) in the treatment group. Although this restriction reduces the sample size, it maintains the original proportion between control and treatment municipalities.

In principle, there could also be an endogeneity problem concerning women’s representation in the council and social spending decisions. In practice, this is not a problem due to institutional restrictions, as voters in Spanish local elections cannot directly show their preferences for male or female politicians by voting for single candidates. The closed lists of a proportional representation system do not give voters any power to affect the order of candidates on party lists. Casas-Arce and Saiz (2015) present evidence for Spanish municipalities suggesting that party leaders were not maximising electoral outcomes prior to the quota system since they included too few women on the lists. According to Cordero et al. (2016), the list order is in the hands of party elites and is more responsive to competing interests within each party than it is to vote maximisation. In addition, the leading left-wing parties have decided to foster female political representation to distinguish themselves from other parties. This triggered a contagion effect across the Spanish party system at the local level, making the lists more homogenous in terms of gender (Kenny and Verge, 2013; Simon and Verge, 2017; Verge, 2020).

When lists look similar, it does not seem likely that voters will choose one party list over the other based on the relative position of female candidates. However, in the unlikely case that voters choose a list according to the order in which it lists female candidates, there still needs to be a sufficient number of voters with the same preferences to translate these preferences into an elected councillor representing this party. Ultimately, the proportion of female councillors results from votes translated into seats according to a proportional representation system, where agreements and coalitions are essential to the outcome. This setup leads to a quasi-exogenous proportion of women in government.

Finally, we study the potential relevance of political budget cycles in the implementation of the reform. According to García and Hayo (2022), budget decisions are linked to the electoral term in the local governments of the Madrid region. Here, we analyse whether the influence of *2015law* is affected by electoral considerations. We investigate this research question by testing whether the reform had a homogeneous effect over the full electoral term.

5.2 Results

First, we check whether the fiscal reform had the intended effect. Table 1 displays the results for Equation (1), our baseline DiD estimation. We report the DiD coefficient of interest only for the standard DiD and the reduced lagged dependent (dynamic DiD) model.[14] Table 1 reports the testing-down restriction from the general model to the reduced model in column 2 as a Chi^2-statistic, which is not significant at any reasonable level of significance. Regarding our variable of interest, we find a negative and statistically significant DiD coefficient, which shows that the reform had the intended effect on treated municipalities, namely lowering social services expenditure by more than 0.9 percentage points (pp). Since average social services spending in our sample is about 5% of the total budget, this implies a reform-induced decrease of about 20%.

Dynamic models allow differentiating between the short- and long-term impacts of explanatory variables. Our focus lies on the short-term effect of the reform, but it is interesting to gauge its long-term impact too, which amounts to a considerable reduction in social spending of 63%. Given our relatively short time dimension, we do not want to emphasise the long-term effect too much, but it certainly suggests that there will be a considerable reduction in social services spending over the years.

Column 1 refers to the standard DiD model and shows that the reform caused a reduction in social services expenditure as a share of total expenditure of more than 2 pp, which amounts to a decrease of 44%. In this simple set-up, short- and long-term impact are not differentiated and the estimated coefficient represents the average treatment effect over the two time perspective. As column 2 shows, the estimated impact of the treatment is only 0.9, which suggests that the short-term reaction of social services spending is only about 1 pp or 20%. However, its long-term impact is much higher, a reduction by more than 3 pp or two-thirds of social services spending. Moreover, standard errors are

[14] See Appendix B1−B4 for the complete set of covariates and all the models.

halved in the dynamic specification, which suggests that, due to standard error decreasing complementarity, it is much more efficiently estimated than the classic DiD specification (Hayo, 2018).

Table 1: Baseline DiD estimation

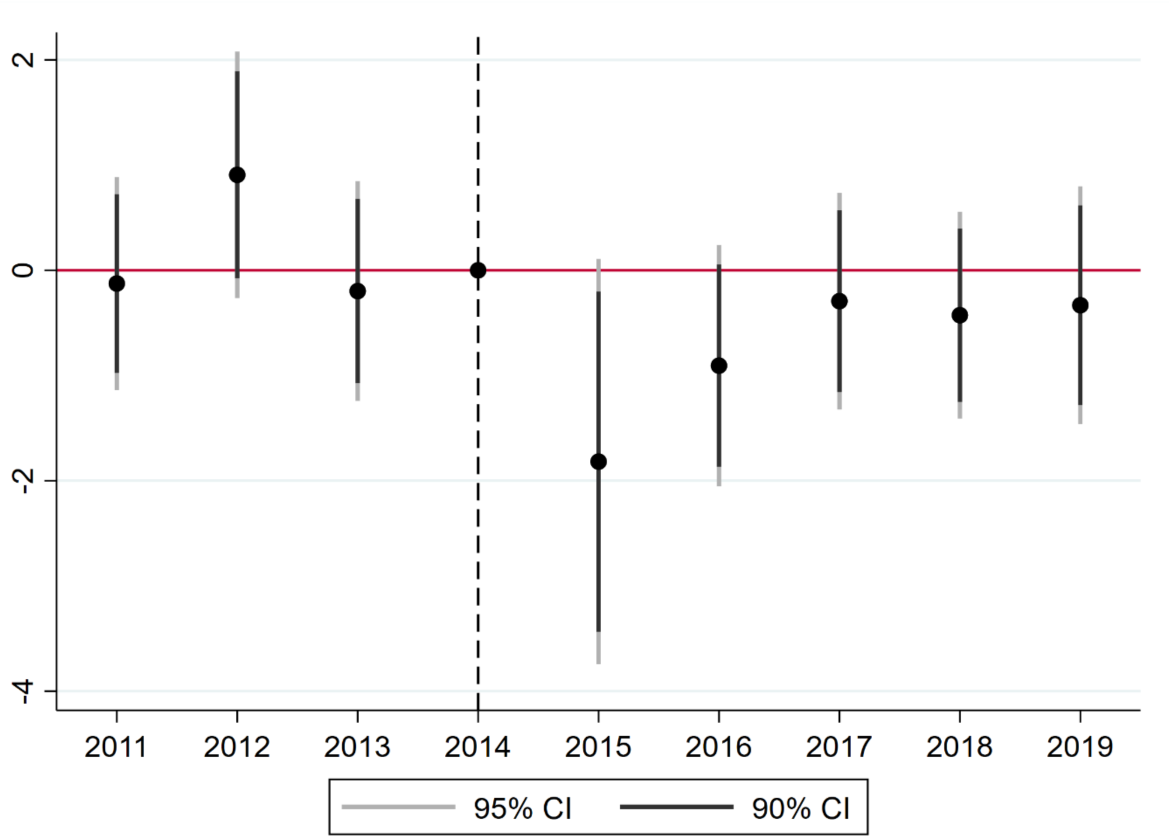
Dependent variable: *Social services spending* (in %)		
	Standard FE	Dynamic FE
	(1)	(2)
*2015law*Treated*	-2.15**	-0.94**
	(0.87)	(0.43)
Municipality FE	Yes	Yes
Time FE	Yes	Yes
Controls	No	Yes
Testing-down restriction	n.a.	$\chi^2(19)= 14.9$
Observations	1,370	1,233
Number of municipalities	137	137

Notes: Values are based on a standard Fixed Effects and the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. When variables with missing observations do not survive the testing-down procedure, in the dynamic estimation, we add available extra observations. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Our event-study estimates for the Extended model are illustrated graphically in Figure 2, which examines the social spending behaviour over time of treated municipalities compared to untreated ones. The plotted estimates for the pre-reform period show no significant differences, which, in combination with Figure 1, indicates that varying pre-trends are not a concern in our study. Focusing on the aftermath of the reform from 2015 onwards shows that the reform had an immediate and strong impact on treated municipalities. For 2015, we find a substantial reduction in social spending of 37% (only significant at the 10% level). The following years also show a reduction in social spending, but the magnitude of the decrease is less than half of that in 2015, and the effect is not statistically significant. This pattern suggests that the reform triggered a fiscal adjustment in 2015, but afterwards, social services expenditure stayed at the post-reform lower value, which underlines the influence of the lagged dependent variable. Based on a model without a lagged dependent variable, Figure D1 in Appendix D shows that all years after the reform are characterised by a significant reduction in social services expenditure, the magnitudes of which are almost equal.

As shown in Table 1 for the baseline model, the results for standard DiD and the specification with a lagged dependent variable and controls are qualitatively the same. Since this conclusion can be generalised, we will only present the estimates for our preferred specification in the main text. The results from the standard DiDs are presented in Appendices C and D.

Figure 2: Baseline event-study DiD estimation



Notes: We estimate the baseline model reported in column 2 of Table 1, 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 in the graph shows the estimated effect, and the bars indicate 95% and 90%, respectively.

As discussed in Section 4, our identification strategy is based on exogenously reforming a budget item that is generally considered particularly interesting to women. The outcome of Table 1 and Figure 2 shows that the reform had the intended effect across all municipalities *on average*. Taking the assumption from the extant literature seriously, namely that women care more about social services spending than men, suggests that female-dominated municipal governments may be more reluctant to decrease spending on this budget item than male-dominated governments.

We consider two gender dimensions: first, we distinguish between male and female mayors and second, we assess the gender influence using the share of female members in the municipal government. To include the possibility of a gender effect along these two dimensions, we interact the DiD term from the previous baseline model in equation (1) with either a female mayor dummy (*Female mayor*) or the share of female politicians in the government (*Female government*).

In Table 2, we report the coefficients of interest for these 3-way interaction models. Column 2 reports a model that interacts the DiD term with a female mayor dummy. Comparing female to male mayors, we find that the triple DiD coefficient is insignificant, indicating no significant differences between untreated and treated municipalities − neither in the short nor the long run − regarding the mayor's sex. Interacting the DiD term with *Female government* yields the model in column 1 of Table 2. Since interpreting point estimates and their significance in the case of continuously interacted variables is problematic (Braumoeller 2004; Brambor et al. 2006), we compute conditional average marginal effects (AMEs) of the DiD coefficient for every value of *Female government*. Figure 3 shows that the gender effect is significantly negative until *Female government* reaches 42%. Thus, treated municipalities spend significantly less on social services until the government becomes gender-balanced as mandated by the Equality Act. The magnitude of the influence of *Female government* is noteworthy. For instance, a government with a share of female politicians of only 28% lowers, on average, social spending by about 2.3 pp, which amounts to a reduction of 47% of average social spending. In contrast, a government with a 42% female share reduces social spending by only one pp, equivalent to 20% of average social spending.

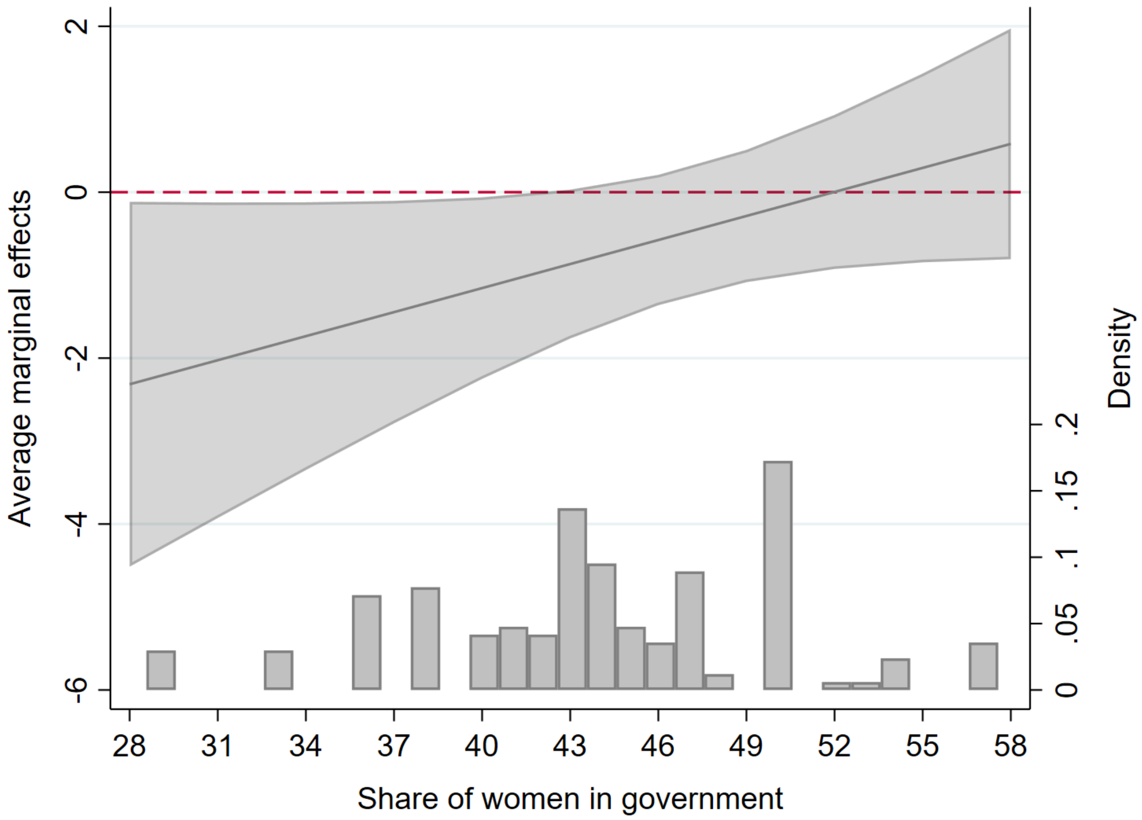
Table 2: DiD estimation interacted with gender variable

Dependent variable: *Social services spending* (in %)		
	% women in government[†] (1)	Female mayor (2)
*2015law*Treated*	-5.02*	-0.77**
	(2.60)	(0.37)
*2015law*Treated*Female government*	0.10*	-
	(0.05)	
*2015law*Treated*Female mayor*	-	-0.53
		(1.76)
Municipality FE	Yes	Yes
Time FE	No	Yes
Controls	Yes	Yes
Testing-down restriction	$\chi^2(19)=15.5$	$\chi^2(20)=21.8$
Observations	1,233	1,233
Number of municipalities	137	137

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. When variables with missing observations do not survive the testing-down procedure, we add available extra observations. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. [†] Time FE cannot be used to compute marginal effects.

Hence, following the 2015 law, there are statistically significant and economically relevant gender differences in the adjustment of social services expenditure. Moreover, these results suggest that the differences between untreated and treated municipalities set out in Table 1 are driven by those governed by male-dominated councils or those barely reaching the lower bounds of a gender-balanced council

Figure 3: AMEs of treated municipalities conditional on the share of females politicians in government



Notes: The point estimates with 95% confidence intervals show the AMEs for treated municipalities after the reform at different values of *Female government*, ranging from 28% (lowest) to 58% (highest). The graphs' bars indicate the density of the variable, *Female government* and are measured on the right-hand y-axis. The underlying regression is reported in Table 2, column 1.

Given that municipalities cannot run deficits, what do gender balanced-governments do to avoid this reduction in social services spending? To answer that question, we look at municipal budgets from two angles. First, we differentiate between expenditures a municipality must provide (mandatory spending) and those it is not required to provide (non-mandatory spending).[15]

In Table A3, we report no statistically significant differences between male-dominated and gender-balanced governments total spending, as well as in mandatory and non-mandatory spending. We examine which specific budget items are cut in a meaningful way by gender-balanced governments and discover that these are, on average, employment and administrative spending. However, male-dominated governments increase expenditures in mobility and road construction. Interestingly, some parts of the political budget cycle literature (Drazen and Eslava, 2010; Repetto, 2018) consider these items to be highly visible expenditures from voters' perspectives.

[15] The classification into mandatory and non-mandatory spending is based on population size, as described on page 8. For a deeper discussion of the topic, see García and Hayo (2022).

So far, we have studied the influence of the two political gender dimensions on the impact of the 2015 budgetary reform independently of each other. However, in municipal-level policymaking, mayors interact intensively with the local council, and it may be the case that the influence of the mayor's sex depends on the share of women in government. To analyse the effect of the mayor's sex conditional on the share of women in the local council, we could employ our baseline estimation interacted with both *Female mayor* and *Female government* at the same time. However, this would imply using 4-way interactions, which are difficult to interpret and inefficiently estimated, especially when, as in our sample, the number of treated units is not very large (Hainmueller et al., 2019; Shieh, 2019). Instead, we adopt a different approach: splitting the sample between municipalities with male-dominated and gender-balanced government and then interacting our DiD term with *Female mayor*. Of our observations, 44% are male-dominated governments, and 56% are gender-balanced governments. Because 19% of all our municipalities belong to the treatment group, the share of treated municipalities of total observations in the subsample of male-dominated governments and gender-balanced governments is 11% and 24%, respectively.

Table 3: DiD estimation on gender-balanced and male-dominated government sample

Dependent variable: *Social services spending* (in %)			
	Gender-balanced		Male-dominated
	Baseline	Female mayor	Baseline
	(1)	(2)	(3)
*2015law*Treated*	-1.06	-0.77	-6.33*
	(0.80)	(0.49)	(3.33)
*2015law*Treated*Female mayor*	-	1.88**	-
		(0.85)	
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Testing-down restriction	$\chi^2(19)=25.8$	$\chi^2(17)=22.1$	$\chi^2(18)=16.3$
Observations	560	452	460
Number of municipalities	80	71	70

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. When variables with missing observations do not survive the testing-down procedure, we add available extra observations. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 3 shows the coefficients of interest for the baseline DiD estimation in the subsample of gender-balanced governments (column 1) and male-dominated governments (column 3), respectively. For the gender-balanced specification, the DiD coefficient is statistically insignificant in the short and long run. In the male-dominated subsample, the reform had the intended effect. The DiD coefficient reports a massive reduction equivalent to 125% of average social spending, but it is only significant at the 10% significance level. Column 2 of Table 3 reports the baseline model interacted with *Female mayor* in the gender-balanced subsample.[16] The estimated effect is significantly positive, implying that in treated municipalities, which are characterised by gender-balanced governments and female mayors, the percentage of social spending increases compared to the combination of gender-balanced councils and male mayors. Tests for the equality of the two coefficients for [*2015law*Treated*] and [*2015law*Treated*Female mayor*] reject the null hypothesis at a significance level of 5%.

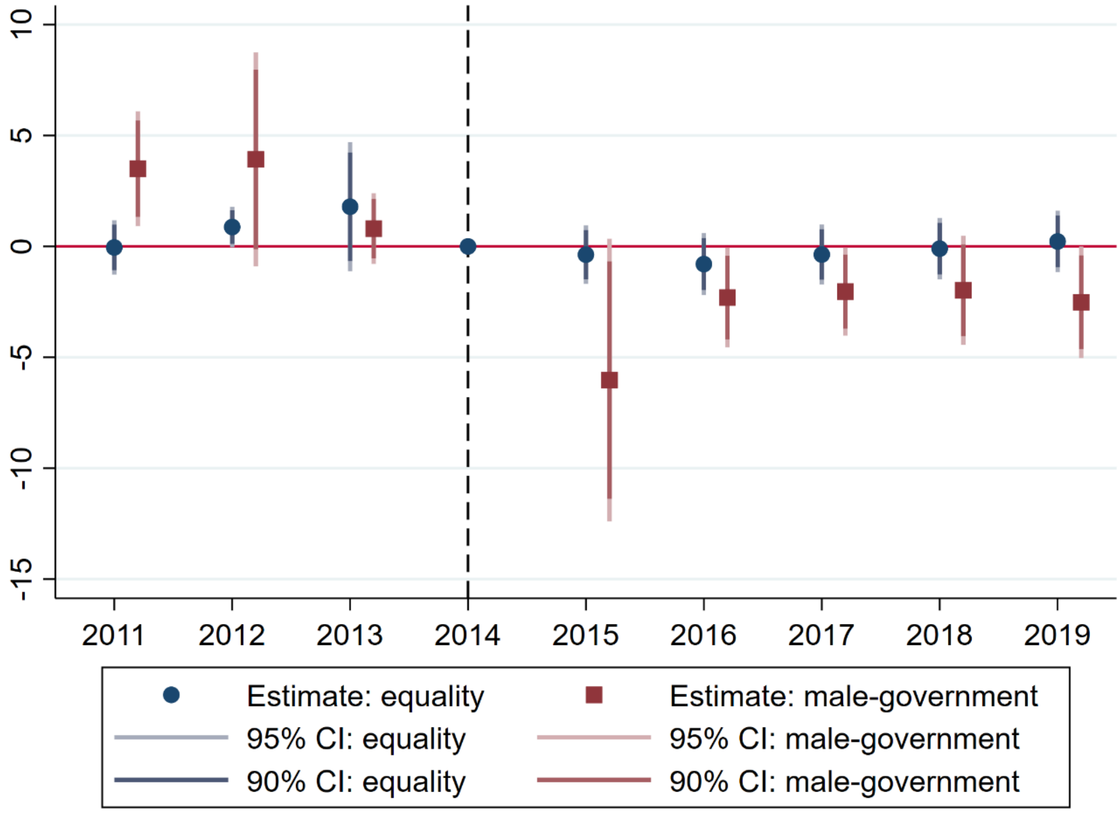
These results corroborate our conjectures (1) and (2), namely that the combination of a female mayor and a gender-balanced government causes the social services item to not decrease in absolute terms and that the decrease in the social services budget item is significantly larger in municipalities with male-dominated governments. Regarding conjecture (1), our results show that the combination of a female mayor and a gender-balanced government leads to social expenditure overshooting, as the social services item not only stays constant but increases in absolute terms. This increase is notable, equivalent to 40% of the average social services spending. This suggests that female-dominated municipal governments' interest in social spending was so great that they completely ignored the intention of the 2015 fiscal reform.

Again, we use event-plot type studies to investigate the dynamics of our DiD estimation for the baseline and the *Female mayor* interaction. In line with the results obtained in Figure 3, the downward adjustment in social services expenditure is driven by the male-dominated government sample (square markers in Figure 4). After the reform, treated municipalities spent significantly less than control municipalities. An exception is 2018; it was a pre-electoral year, suggesting that social spending was somewhat affected by PBC-related budgeting activities.

We also perform an event study analysis for the interaction of treated municipalities with *Female mayor* in the subsample of gender-balanced government (Figure 5). We find significant differences between male and female mayors in 2017, 2018, and 2019. The triple DiD coefficient is positive, which indicates that − after the reform and conditional on gender-balanced governments − female mayors tended to spend more on social services than their male counterparts. Generalising this result, it appears that a couple of years after implementing such a fiscal reform, mayors are willing and able to reassert their preferences concerning specific budget items.

[16] In the subsample of male-dominated governments, the number of municipalities with a female mayor after the reform is too small for estimation.

Figure 4: Baseline event-study DiD estimation in both subsamples



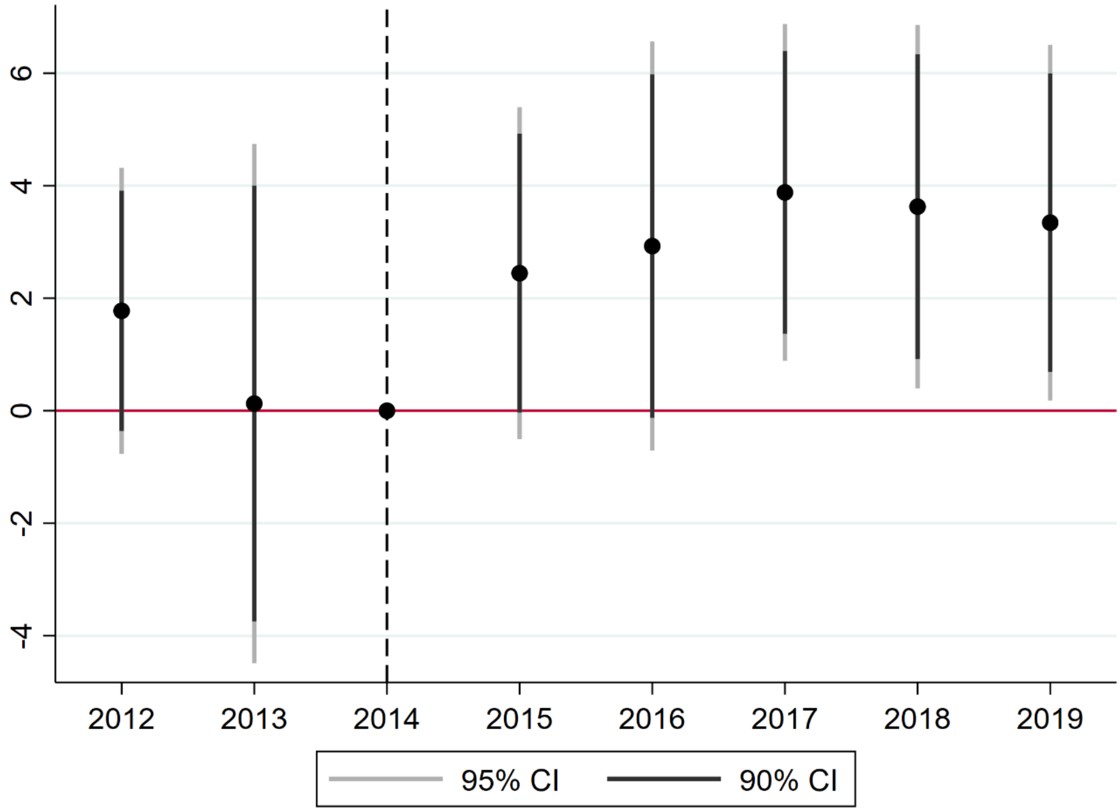
Notes: We estimate the baseline models reported in Table 3, columns 1 and 3, but *2015law* is substituted by year dummies. The effect is set to 0 in the last year before the reform, and the resulting reference year 2014 is marked by a dashed vertical line. Each dot (and square) in the graph shows the estimated effect, and the bars indicate 95% and 90% confidence intervals, respectively.

To analyse the potential influence of PBCs, we examine the effect of the reform conditional on the electoral term. Table 4 reports the estimated effect of interacting our baseline model with the pre-electoral dummy on the entire sample and on the gender-balanced and male-dominated government subsamples. The first estimates in columns 1 of Table 4 reflect our previous result: treated municipalities significantly reduced their share of social spending in the aftermath of the 2015 reform. However, the sign of the effect is reversed for the interaction term with pre-electoral years, which implies that treated municipalities tend to increase their share of social services expenditure before elections.

The magnitudes of the two estimated effects are very similar in their absolute values, suggesting a transfer of social services funds from non-pre-electoral to pre-electoral years. Indeed, a t-test of the equality of both coefficients rejects the null hypothesis at the 5% significance level, whereas we cannot reject that both coefficients are equal in absolute terms. In the gender-balanced subsample, column

2, there is a weakly significant increase in social services spending during pre-electoral years. Again, the equality of both coefficients in absolute terms cannot be rejected at the 5% significance level.

Figure 5: Event-study DiD estimation with female mayor interaction



Notes: We estimate the baseline model interacted with Female mayor reported in Table 3, column 2, but variable 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 (and square) in the graph shows the estimated effect, and the bars indicate 95% and 90%, respectively.

For the male-dominated subsample, in column 3 of Table 4, the reform had a significant negative effect during non-pre-electoral years, which, however, is partially offset by a hike in social services spending in pre-electoral years. Here, the coefficients are not similar in magnitude, which indicates that, overall, there is a reduction in social services funding. At the 5% significance level, we cannot reject that both coefficients are equal in absolute terms, which means we cannot exclude the possibility that the hike in social services spending in pre-election years exactly offsets the reduction during non-election years.

Thus, our results for the specific case of the 2015 reform are in line with the general findings reported by García and Hayo (2022) on PBCs in the Madrid region: under specific circumstances, such as a specific preference for social services, female politicians can behave as opportunistically motivated as male politicians.

Table 4: DiD estimation conditional on the electoral term

Dependent variable: *Social services spending* (in %)			
	Full sample	Gender-balanced	Male-dominated
	(1)	(2)	(3)
*2015law*Treated*	-1.20**	-1.42	-7.24**
	(0.54)	(0.93)	(3.50)
*2015law*Treated*Pre-election*	1.10*	1.32*	3.31***
	(0.60)	(0.69)	(1.23)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Testing-down restriction	$\chi^2(19)=15.3$	$\chi^2(19)=26.2$	$\chi^2(18)=19.3$
Observations	1,233	560	460
Number of municipalities	137	80	70

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. When variables with missing observations do not survive the testing-down procedure, we add available extra observations. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

6. Robustness

In Section 5.1, we argued that it is unlikely that the mayor's gender is endogenous in a restricted sample of municipalities where a female (or male) mayor was elected in both the 2011 and 2015 elections. However, to further address any remaining doubts about the validity of the analysis, we narrow our focus to the subsample of closely contested elections and apply a regression discontinuity design. The rationale underlying this approach is that elections won by a sufficiently narrow margin are very similar to elections lost by a narrow margin. Therefore, whether a mayor is male or female is essentially a matter of chance. We adopt a common practice in the literature by examining mixed-gender races involving the two most voted parties (Gagliarducci and Paserman, 2012; Accettura and Profeta, 2022; Carozzi and Gago, 2023).[17]

By concentrating on municipalities where elections were decided by a margin of 5 pp or less, Table 5 replicates Table 2. The choice of the 5 pp bandwidth represents a compromise between the 'optimal' bandwidth determined by the procedure proposed by Calonico et al. (2014) and a margin size that can reasonably be considered as indicative of a 'close-election'.[18]

[17] With the regression discontinuity method for closed-list proportional representation developed by Curto-Grau et al. (2018), our numerical estimator did not converge when the bandwidth was restricted to a plausible range to be considered close-election.

[18] The 'optimal' bandwidth tends to be larger than 5%. However, our results are robust to the choice of bandwidth (available on request).

Table 6: DiD estimation interacted with gender variable in a mixed-gender close election sample

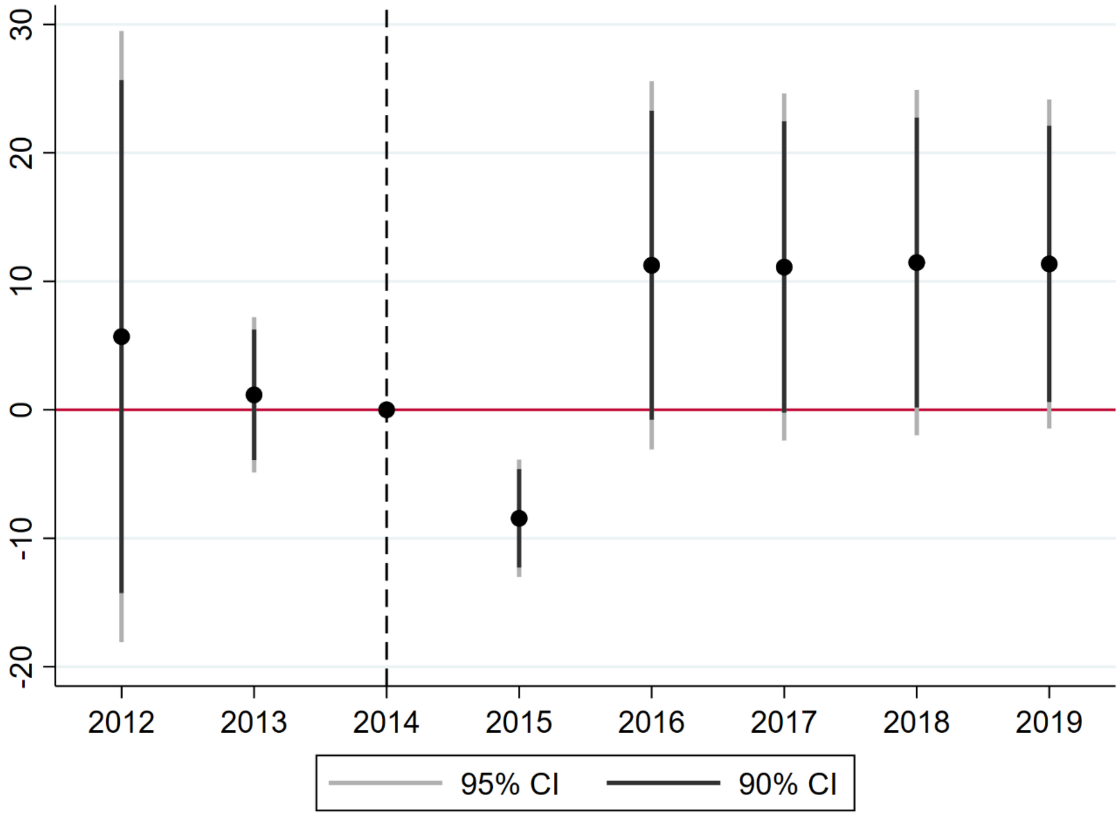
Dependent variable: *Social services spending* (in %)		
	% women in government	Female mayor
	(1)	(2)
*2015law*Treated*	-127.92***	0.73**
	(24.30)	(0.96)
*2015law*Treated*Female government*	3.26***	-
	(0.59)	
*2015law*Treated*Female mayor*	-	-9.99***
		(2.49)
Municipality FE	Yes	Yes
Time FE	No	Yes
Controls	Yes	Yes
Observations	184	184
Bandwidth (percent)	5.00	5.00

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the difference in votes between the two most voted party lists is equal or less than 5%. Cluster-robust standard errors are used (clusters: number of municipalities). *, **, and *** indicate significance at a 10%, 5%, and 1% level, respectively.

The estimates in column 1 are in line with the ones reported in Table 2, but the coefficients are more statistically significant and larger in magnitude. In contrast, column 2 indicates that female mayors reduce social expenditures relatively more than their male counterparts. This effect is in stark contrast to the effect observed in column 2, Table 3, where female mayors in gender-balanced governments increase social spending compared to male ones. Gagliarducci and Paserman (2012) note that women in office often encounter more challenges, particularly when councils are male dominated. This helps explain why a highly competitive close-election setting may influence the behaviour of female mayors differently than situation where victories are clear.

In Figure 6, we employ event-plot type studies to examine the dynamics of our DiD estimation for the *Female mayor* interaction. Consistent with the results presented in Figure 2, the reduction in social services expenditure is observed only in 2015. However, female mayors exhibit a higher increase in social spending compared to their male counterparts in 2018 and 2019. This finding aligns with the results shown in Figure 5.

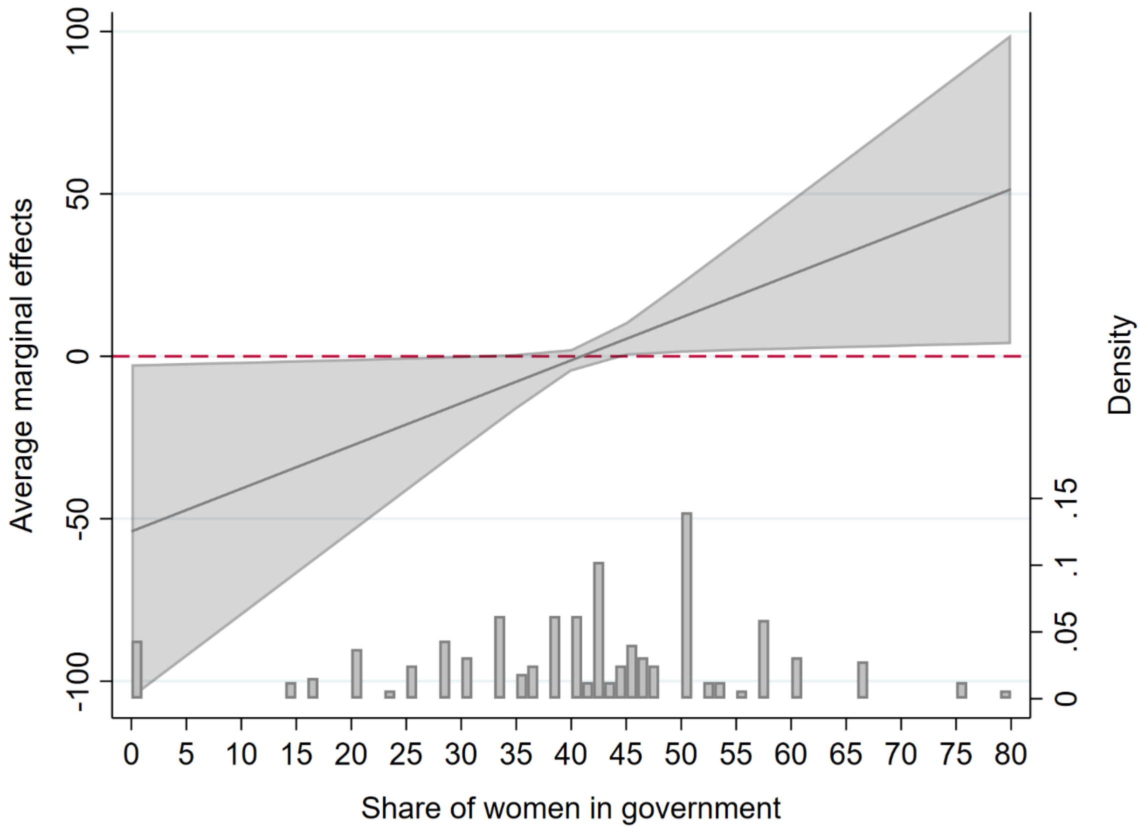
Figure 6: Event-study DiD estimation with female mayor interaction in a mixed-gender close election sample



Notes: We estimate the model reported in Table 6, column 2, but variable 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 (and square) in the graph shows the estimated effect, and the bars indicate 95% and 90%, respectively.

Turning to the representation of women in government, Figure 7 illustrates a negative gender effect in more male-dominated governments. When the share of women in government increases to more than 45%, then the more female-dominated governments show a significant increase in social service spending. This result mirrors the pattern we observed in Figure 3 and D2, where male-dominated governments appear to cause a reduction in social spending. Given the challenges faced by female leaders, their spending behaviour appears to be contingent on the proportion of female politicians in government. This suggests that a sufficiently large group of women in government is needed before female mayors can effectively implement their spending preferences.

Figure 7: AMEs of treated municipalities conditional on the share of females politicians in government in a mixed-gender close election sample



Notes: The point estimates with 95% confidence intervals show the AMEs for treated municipalities after the reform at different values of *Female government*, ranging from 0% (lowest) to 80% (highest). The graphs' bars indicate the density of the variable, *Female government* and are measured on the right-hand y-axis. AMEs are obtained on a standard Fixed Effects estimation with a 7% bandwidth.

7. Conclusion

Using a dataset on Spanish municipalities, we investigate gender differences in the fiscal behaviour of politicians when confronted with an exogenous change in the institutional framework. In our analysis, we exploit the introduction of a budget modification aimed at substantially reducing the scope of social services at the municipal level. A widespread assumption in the literature is that women hold distinctive social preferences and favour specific fiscal budget items, particularly social, health, and education spending (Hessami and da Fonseca, 2020; Bagues and Campa, 2021). In light of this discussion, our general hypothesis is that conditional on politicians' gender, a reform on an item traditionally considered to be of particular interest to women will have a different impact on affected municipalities. Specifically, we conjecture that the combination of a female mayor and a gender-balanced government causes the social services budget item to react (1) very little, if at all, in absolute terms and (2) certainly less than the adjustment taking place in more male-dominated districts.

In December 2013, the 'Local Government Rationalisation and Sustainability Act' (No. 27/2013) was introduced in Spain. Before the change in the law, municipalities with more than 20,000 inhabitants had the responsibility of managing the provision of social services. The reform limits this budget function to identifying situations where social assistance is needed. To study the effect of this fiscal reform, we use data from the Madrid region (Comunidad Autónoma de Madrid) from 2010−2019. Using a DiD estimator, we compare gender differences in social services spending before and after the reform between municipalities below 20,000 inhabitants (control group) and municipalities with 20,000 or more inhabitants (treatment group).

First, we compare social services spending before and after the reform without considering gender aspects. Our DiD estimates show that the reform had the intended effect across all municipalities by reducing social spending by about 20%. The reform had an immediate and substantial effect in 2015, whereas the following years showed a more minor and non-significant reduction in social spending.

Second, when conditioning the analysis on politicians' sex, a more differentiated picture emerges. On the one hand, as long as the governments are not gender-balanced as mandated by the Equality Act, treated municipalities spend significantly less on social services after the reform. The effect ranges from an almost 50% reduction of average social spending for governments with a less than 30% share of female politicians to a 20% reduction for governments with a 42% share of female politicians. On the other hand, there are no significant differences between untreated and treated municipalities when conditioning on the mayor's sex. However, in a regression discontinuity analysis conducted on a subsample, we find that female mayors reduce social expenditures relatively more than their male counterparts. This result suggests that in a highly competitive close-election setting, female mayors may behave differently than in situation where victories are clear. However, this conclusion rests on a small set of observations and may not be particularly robust.

Third, the sample is split between male-dominated and gender-balanced governments, and the only reduction is found in the male-dominated government sample, corroborating our conjecture (2). When the effect is conditioned on the mayor's sex, we discover that gender-balanced governments combined with female mayors increase social services spending by 40%, which supports conjecture (1). Moreover, this increase in social services spending is significantly different from that of gender-balanced governments combined with male mayors, which reduces this type of expenditure in line with our conjecture (2).

Our analysis reveals that a reform that had the intended effect, on average, can trigger different responses conditional on politicians' sex. Although social spending was significantly reduced after the introduction of the reform, this adjustment is mainly driven by male-dominated governments. Despite being widely used to study gender differences in the behaviour of politicians, in our study, the mayor's sex is only relevant in the subsample of gender-balanced governments or when we focus on close elections. At the municipal level, mayors interact intensively with the local council, and focusing only on the mayor likely leads to an incomplete picture of gender differences in municipal-level policymaking. Our findings support the claim that social spending is, on average, of particular importance to female politicians, as they appear to go as far as bending the law to uphold their interests.

In addition, we analyse the effect of the reform conditional on the electoral term. We still find that, on average, treated municipalities significantly reduced the share of social spending in the aftermath of the 2015 reform. However, the effect is reversed in pre-electoral years, which implies that treated municipalities tend to increase their share of social spending before elections compared to non-election years. Since this effect seems to be mainly due to decisions made by local governments dominated by women, we provide evidence that female politicians choose to engage in PBC-related activities under specific circumstances, such as a preference for social spending. Thus, in contrast to some findings in the literature (Brollo and Troiano, 2016; Accettura and Profeta, 2022), our results suggest that female politicians can behave as opportunistically motivated as their male counterparts.

Regarding the limits of our dataset, we would like to note that compared to the rest of the country, our sample region under-represents small municipalities and over-represents large ones. While this over-representation of large municipalities provides us with a greater number of treated municipalities, it might also endanger the external reliability (often called 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'. Due to the Madrid region's highly homogenous set of budget rules and the fulfilled requirements for valid difference-in-differences estimation, our study offers a high degree of internal validity.

To summarise, contrary to studies reporting that an exogenous increase in the share of female politicians has no impact on the composition of public spending (Geys and Sørensen, 2019; Bagues and Campa, 2021), we show that higher female representation can generate notable variations in public municipal spending priorities, at least once institutional and budgetary constraints are taken into consideration. More generally, we provide evidence that female leaders are not necessarily more law-compliant than their male counterparts if the law opposes their interests.

References

Accettura, C. and Profeta, P. (2022). Gender differences in political budget cycles. Paper presented at the European Public Choice Society conference in Braga, April.

Aidt, T.S. and Dallal, B. (2008). Female voting power: The contribution of women's suffrage to the growth of social spending in Western Europe (1869—1960). *Public Choice* 134, 391–417.

Aidt, T.S., Dutta, J., and Loukoianova, E. (2006). Democracy comes to Europe: Franchise extension and fiscal outcomes 1830—1938. *European Economic Review* 50, 249–283.

Alesina, A. and La Ferrara, E. (2005). Preferences for redistribution in the land of opportunities. *Journal of Public Economics* 89, 897–931.

Andersen, S., Bulte, E., Gneezy, U., and List, J. A. (2008). Do women supply more public goods than men? Preliminary experimental evidence from matrilineal and patriarchal societies. *American Economic Review* 98, 376–381.

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.

Baltrunaite, A., Bello, P., Casarico, A., and Profeta, P. (2014). Gender quotas and the quality of politicians. *Journal of Public Economics* 118, 62–74.

Baskaran, T. and Hessami, Z. (2019). Competitively elected women as policy makers. *CESifo Working Paper* No. 8005.

Bertocchi, G. (2011). The enfranchisement of women and the welfare state. *European Economic Review* 55, 535–553.

Besley, T., Folke, O., Persson, T., and Rickne, J. (2017). Gender quotas and the crisis of the mediocre man: theory and evidence from Sweden. *American Economic Review* 107, 2204–2242.

Bhalotra, S. and Clots-Figueras, I. (2014). Health and the political agency of women. *American Economic Journal: Economic Policy* 6, 164–197.

Braga, M. and Scervini, F. (2017). The performance of politicians: The effect of gender quotas. *European Journal of Political Economy* 46, 1–17.

Brambor, T., Clark, W.R., and Golder, M. (2006). Understanding interaction models: Improving empirical analyses. *Political Analysis* 14, 63–82.

Braumoeller, B.F. (2004). Hypothesis testing and multiplicative interaction terms. *International Organization* 58, 807–820.

Breitung, J., Kripfganz, S., and Hayakawa, K. (2022). Bias-corrected methods of moment estimators for dynamic panel data models. *Econometrics and Statistics* (in press).

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.

Cabaleiro-Casal, R. and Buch-Gómez, E. (2018). Adjustments in municipal fiscal crises. Are they different according to the gender of the mayor? *Local Government Studies* 44, 255–274.

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.

Carozzi, F. and Gago, A. (2023). Who promotes gender-sensitive policies?. *Journal of Economic Behavior & Organization* 206, 371–405.

Casarico, A., Lattanzio, S., and Profeta, P. (2022). Women and local public finance. *European Journal of Political Economy* 72, 102096.

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

Cerreda, M.A. (2015). El impacto de la ley 27/2013, de 27 de diciembre, de racionalización y sostenibilidad de la Administración local, en la distribución de competencias en materia de educación, salud, sanidad y servicios sociales. *Revista D'estudis Autonòmics i Federals* 22, 219–263.

Chattopadhyay, R. and Duflo, E (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica* 72, 1409–1443.

Clayton, A. and Zetterberg, P. (2018). Quota shock: Electoral gender quotas and government spending priorities worldwide. *Journal of Politics* 80, 916–932.

Clots-Figueras, I. (2011). Women in politics: Evidence from the Indian states. *Journal of Public Economics* 95, 664–690.

Clots-Figueras, I. (2012). Are female leaders good for education? Evidence from India. *American Economic Journal: Applied Economics* 4, 212–244.

Cordero, G. Jaime-Castillo, A.M., and Coller, X. (2016). Candidate selection in a multilevel state. *American Behavioral Scientist* 60, 853–868.

Corneo, G. and Grüner, H. P. (2002). Individual preferences for political redistribution. *Journal of Public Economics* 83, 83–107.

De Paola, M., Scoppa, V., and Lombardo, R. (2010). Can gender quotas break down negative stereotypes? Evidence from changes in electoral rules. *Journal of Public Economics* 94, 344–353.

De Paola, M., Scoppa, V., and De Benedetto, M. A. (2014). The impact of gender quotas on electoral participation: Evidence from Italian municipalities. *European Journal of Political Economy* 35, 141–157.

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

Esteve-Volart, B. and Bagues, M. (2012). Are women pawns in the political game? Evidence from elections to the Spanish senate. *Journal of Public Economics* 96, 387–399.

Ferreira, F. and Gyourko, J. (2014). Does gender matter for political leadership? The case of U.S. mayors. *Journal of Public Economics* 112, 24–39.

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.

Funk, K. D. and Philips, A. Q. (2019). Representative budgeting: Women mayors and the composition of spending in local governments. *Political Research Quarterly* 72, 19–33.

Funk, P. and Gathmann, C. (2015). Gender gaps in policy making: Evidence from direct democracy in Switzerland. *Economic Policy* 30, 141–181.

Gagliarducci, S. and Paserman, M. D. (2012). Gender interactions within hierarchies: Evidence from the political arena. *Review of Economic Studies* 79, 1021–1052.

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.

Geys, B. and Sørensen, R.J. (2019). The impact of women above the glass ceiling: Evidence from a Norwegian executive gender quota reform. *Electoral Studies* 60.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica* 77, 1637–1664.

Hainmueller, J., Mummolo, J., and Xu, Y. (2019). How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice. *Political Analysis* 27, 163–192.

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.

Hendry, D. F. (1993). *Econometrics: alchemy or science? Essays in econometric methodology*. Blackwell Publishers, Oxford.

Hernández-Nicolás, C.M., Martín-Ugedo, J. F., and Mínguez-Vera, A. (2018). Women mayors and management of Spanish councils: An empirical analysis. *Feminist Economics* 24, 168–191.

Hessami, Z. and da Fonseca, M.L. (2020). Female political representation and substantive effects on policies: A literature review. *European Journal of Political Economy* 63, 101896.

Kenny, M. and Verge, T. (2013). Decentralization, political parties, and women's representation: Evidence from Spain and Britain. *Journal of Federalism* 43, 109–128.

Kiviet, J.F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of Econometrics* 68, 53–78.

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

Lippmann, Q. (2022). Gender and lawmaking in times of quotas. *Journal of Public Economics* 207, 104610.

Lott, J. R. and Kenny, W. (1999). Did women's suffrage change the size and scope of government? *Journal of Political Economy* 107, 1163–1198.

Navarro-Galera, A., Buendía-Carrillo, D., Lara-Rubio, J., and Rayo-Cantón, S. (2017). Do political factors affect the risk of local government default? Recent evidence from Spain. *Lex Localis—Journal of Local Self-Government* 15, 43–66.

Nickell, S. (1981). Biases in dynamic models with fixed effects. *Econometrica* 49, 1417–1426.

O'Brien, D. Z. and Rickne, J. (2016). Gender quotas and women's political leadership. *American Political Science Review* 110, 112–126.

Oswald, A. J. and Powdthavee, N. (2010). Daughters and left-wing voting. *Review of Economics and Statistics* 92, 213–227.

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

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

Shieh, G. (2019). Effect size, statistical power, and sample size for assessing interactions between categorical and continuous variables. *British Journal of Mathematical and Statistical Psychology* 72, 136–154.

Simon, P. and Verge, T. (2017). Gender quotas and political representation in Spain and Portugal: Electoral competition, learning and emulation. *South European Society and Politics* 22, 179–195.

Slegten, C. and Heyndels, B. (2019). Within-party sex gaps in expenditure preferences among Flemish local politicians. *Politics & Gender* 16, 768–791.

Sweeting, D. (2009). The institutions of 'strong' local political leadership in Spain. *Environment and Planning C: Government and Policy* 27, 698–712.

Verge, T. (2020). Political party gender action plans: Pushing gender change forward beyond quotas. *Party Politics* 26, 238–248.

Washington, E. L. (2008). Female socialization: How daughters affect their legislator fathers' voting on women's issues. *American Economic Review* 98, 311–332.

Weeks, A. C. (2019). Quotas and party priorities: Direct and indirect effects of quota laws. *Political Research Quarterly* 72, 849–862.

Appendix A

Table A1: Descriptive statistics

Variable	*Description*	*Min*	*Max*	*Mean*	*Std. Dev.*
Social services spending	Expenditures on social services in relation to total expenditures (in %).	0	72.28	4.94	9.33
Own revenues	Revenues from direct taxes, indirect taxes, and fees in relation to total revenues (in %).	2.63	88.26	55.56	16.81
Current transfers	Current transfers from higher-level government (regional and central) in relation to total revenues (in %).	2.79	82	31.22	12.17
Capital transfers	Capital transfers from higher-level government (regional and central) in relation to total revenues (in %).	-0.19	70.31	6.76	12.05
Debt	Municipal debt in relation to total expenditures (in %).	0	940.06	39.80	71.47
Rent	Municipal income in relation to total expenditures (in %).	73.98	3600.17	1371.99	628.18
Unemployment rate of youth	Percentage of registered unemployed below 25 in relation to labour force.	0	20	3.19	2.52
Unemployment rate of adults	Percentage of registered unemployed between 25 and 45 in relation to labour force.	0	81.78	29.24	15.57
Unemployment rate of elderly	Percentage of registered unemployed above 45 in relation to labour force.	0	40.48	13.34	7.4
Population density	Inhabitants per km^2.	1.06	7635.99	513.87	1121.98
Share of youth	Share of the population below 15 in relation to total population (in %).	0	36.48	26.15	5.35
Share of retired	Share of the population above 65 in relation to total population (in %).	3.4	54.1	15.68	6.78
Share of immigration	Share of non-Spanish inhabitants in relation to total population (in %).	0	40.42	13.15	5.64
Female mayor	Dummy variable taking value 1 when mayor is female (0 otherwise).	0	1	0.23	0.42
Mayor's age	Age of the mayor in years.	25	76	49.69	9.26
Mayor's primary education	Dummy variable taking value 1 when the mayor obtained primary education (0 otherwise).	0	1	0.01	0.12
Mayor's secondary education	Dummy variable taking value 1 when the mayor obtained high school or similar education (0 otherwise).	0	1	0.36	0.48
Mayor's higher education	Dummy variable taking value 1 when the mayor obtained university education (0 otherwise).	0	1	0.39	0.49

Left	Dummy variable taking value 1 when the governing body has left-wing ideology (0 otherwise).	0	1	0.23	0.42
Right	Dummy variable taking value 1 when the governing body has right-wing ideology (0 otherwise).	0	1	0.64	0.48
Parties in government	Number of parties in the government.	1	7	1.81	1.31
Female government	Share of women in the governing body in relation to total councillors in the governing body (in %).	0	100	37.96	16.79
2015 law	Dummy variable taking the value 1 from 2015 onwards (0 otherwise).	0	1	0.50	0.50
Treated	Dummy variable that takes the value 1 for municipalities with 20,000 or more inhabitants (0 otherwise).	0	1	0.19	0.39
Pre-election	Dummy variable taking the value 1 in the year preceding a local election (0 otherwise).	0	1	0.30	0.46
Election	Dummy variable taking the value 1 in the year of a local election (0 otherwise).	0	1	0.30	0.46
Post-election	Dummy variable taking the value 1 in the year after a local election (0 otherwise).	0	1	0.20	0.40

Table A2: Covariates before and after the 2015 fiscal reform

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Pre-			Post-		DiD
	Treatment	Control	Difference	Treatment	Control	Difference	(3)-(6)
Total expenditures per capita	1083.78	1665.60	-581.82***	950.20	1578.82	-628.62***	46.80
	(23.42)	(57.56)	(121.82)	(14.30)	(57.54)	(117.88)	(169.49)
Own revenues	65.21	50.89	14.32***	68.03	55.00	13.03***	1.29
	(0.88)	(0.69)	(1.50)	(0.64)	(0.59)	(1.24)	(1.94)
Current transfers	25.45	33.39	-7.94***	26.37	31.51	-5.14***	-2.80*
	(0.60)	(0.52)	(1.14)	(0.59)	(0.43)	(0.92)	(1.46)
Capital transfers	0.59	8.25	-7.66***	0.68	8.21	-7.52***	-0.14
	(0.14)	(0.49)	(1.03)	(0.13)	(0.49)	(0.99)	(1.43)
Debt	66.07	30.29	35.78***	76.04	33.77	42.27***	-6.48
	(4.91)	(1.73)	(4.33)	(7.45)	(3.11)	(7.33)	(8.55)
Rent	1627.68	1186.52	441.16***	1990.72	1353.57	637.15***	-195.99**
	(34.65)	(21.53)	(48.26)	(36.93)	(24.21)	(52.69)	(71.24)
Unemployment rate of youth	3.28	4.44	-1.16***	1.63	2.30	-0.68***	-0.48*
	(0.14)	(0.11)	(0.24)	(0.08)	(0.07)	(0.14)	(0.28)
Unemployment rate of adults	27.52	34.57	-7.05***	18.74	26.98	-8.24***	1.19
	(0.92)	(0.61)	(1.35)	(0.77)	(0.56)	(1.21)	(1.81)
Unemployment rate of elderly	11.03	13.74	-2.71***	9.52	14.48	-4.96***	2.25**
	(0.37)	(0.27)	(0.60)	(0.36)	(0.31)	(0.65)	(0.89)
Population density	2001.47	138.20	1863.27***	1964.70	139.24	1825.46***	37.81
	(142.57)	(7.19)	(69.22)	(132.81)	(7.44)	(35.90)	(96.05)
Share of youth	28.93	25.93	2.99***	28.74	25.17	3.57***	-0.58
	(0.23)	(0.20)	(0.44)	(0.25)	(0.22)	(0.46)	(0.64)
Share of retired	10.60	15.73	-5.12***	13.06	17.36	-4.30***	-0.82
	(0.25)	(0.27)	(0.58)	(0.30)	(0.26)	(0.55)	(0.79)
Share of immigration	14.69	14.64	0.27	12.06	11.48	0.58	-0.31
	(0.40)	(0.23)	(0.52)	(0.30)	(0.19)	(0.42)	(0.67)
Female mayor	0.15	0.25	-0.10***	0.21	0.25	-0.04	-0.06
	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)	(0.04)	(0.05)
Mayor’s age	49.91	48.97	0.93	49.06	50.35	-1.29	2.22*

	(0.73)	(0.36)	(0.82)	(0.82)	(0.39)	(0.89)	(1.21)
Mayor’s primary education	0.00	0.03	-0.03*	0.00	0.01	-0.01	-0.03*
	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)
Mayor’s secondary education	0.16	0.43	-0.27***	0.17	0.39	-0.23***	0.04
	(0.03)	(0.02)	(0.04)	(0.03)	(0.02)	(0.04)	(0.06)
Mayor’s higher education	0.62	0.32	0.30***	0.64	0.33	0.31***	0.02
	(0.04)	(0.02)	(0.04)	(0.04)	(0.02)	(0.04)	(0.06)
Female government	43.35	34.10	9.25***	43.96	39.05	4.90***	4.35**
	(0.54)	(0.67)	(1.43)	(0.46)	(0.69)	(1.43)	(2.02)
Left	0.22	0.18	0.05	0.36	0.24	0.12***	-0.07
	(0.03)	(0.02)	(0.03)	(0.04)	(0.02)	(0.04)	(0.05)
Right	0.76	0.71	0.05	0.56	0.57	-0.01	0.06
	(0.03)	(0.02)	(0.04)	(0.04)	(0.02)	(0.04)	(0.06)
Parties in government	1.68	1.51	0.17*	3.18	1.82	1.36***	-1.19***
	(0.09)	(0.04)	(0.09)	(0.14)	(0.05)	(0.12)	(0.15)

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

Table A3: Expenditures in male-dominated and gender-balanced governments before and after the 2015 fiscal reform

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		Gender-equality			Male-dominated		Difference
	Pre-	Post-	Difference	Pre-	Post-	Difference	(3)-(6)
Total expenditures per capita	1098.38	954.38	144.00***	1048.69	934.32	114.37*	29.64
	(26.60)	(16.84)	(30.57)	(47.68)	(25.15)	(59.42)	(63.29)
Mandatory spending	39.92	40.25	-0.33	40.36	41.15	-0.79	0.46
	(0.80)	(0.77)	(1.11)	(0.82)	(1.18)	(1.39)	(2.10)
Non-mandatory spending	2.81	1.54	1.27	1.28	1.37	-0.09	1.36
	(0.75)	(0.35)	(0.82)	(0.25)	(0.28)	(0.92)	(1.31)
Mobility	0.67	0.67	-0.00	0.38	1.04	-0.66**	0.66*
	(0.14)	(0.12)	(0.19)	(0.07)	(0.35)	(0.31)	(0.37)
Housing	0.26	0.16	0.10	0.06	0.05	0.01	0.09
	(0.07)	(0.20)	(0.07)	(0.03)	(0.02)	(0.04)	(0.12)
Road construction	3.32	3.29	0.03	2.47	4.38	-1.91**	1.94**
	(0.33)	(0.27)	(0.42)	(0.35)	(0.48)	(0.00)	(0.80)
Education	5.96	6.15	-0.19	5.63	5.46	-0.17	-0.36
	(0.21)	(0.29)	(0.29)	(0.30)	(0.34)	(0.45)	(0.57)
Health	0.17	0.25	-0.08	0.27	0.44	-0.17	0.09
	(0.04)	(0.04)	(0.06)	(0.07)	(0.12)	(0.14)	(0.13)
Employment	1.63	1.18	0.45**	1.78	0.88	0.90**	0.45
	(0.12)	(0.07)	(0.14)	(0.29)	(0.02)	(0.35)	(0.31)
Administration	2.90	1.86	1.04**	1.35	1.25	0.09	0.95
	(0.35)	(0.28)	(0.45)	(0.41)	(0.35)	(0.56)	(0.84)
Environment	5.79	5.63	0.16	5.10	6.05	-0.95**	1.12
	(0.32)	(0.20)	(0.36)	(0.27)	(0.36)	(0.44)	(0.68)

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

Appendix B: General and reduced models with full set of covariates

Table B1: Baseline DiD estimation

Dependent variable: *Social services spending* (in %)			
	General	Reduced	
		Smaller sample	Larger sample
	(1)	(2)	(3)
Lag social services	0.72***	0.73***	0.70***
	(0.12)	(0.13)	(0.13)
Own revenues	0.00	-	-
	(0.01)		
Lag own revenues	0.01	-	-
	(0.02)		
Current transfers	-0.00	-	-
	(0.02)		
Lag current transfers	0.03	-	-
	(0.04)		
Capital transfers	-0.01	-	-
	(0.02)		
Lag capital transfers	0.03*	-	-
	(0.02)		
Lag debt	0.00	-	-
	(0.00)		
Lag rent	-0.00	-	-
	(0.00)		
Lag unemployment rate of youth	0.10	-	-
	(0.11)		
Lag unemployment rate of adults	0.00	-	-
	(0.03)		
Lag unemployment rate of elderly	-0.02	-	-
	(0.04)		
Lag population density	-0.00	-	-
	(0.00)		
Lag Share of youth	-0.05	-	-
	(0.17)		
Lag Share of retired	0.15	-	-
	(0.09)		
Lag Share of immigration	-0.16*	-0.21**	-0.20**
	(0.09)	(0.10)	(0.09)
Mayor’s age	0.01	-	-
	(0.01)		
Mayor’s secondary education	-0.01	-	-
	(0.29)		
Mayor’s higher education	-0.27	-	-
	(0.28)		
Female mayor	0.16	0.12	-0.19
	(0.49)	(0.51)	(0.52)

Female government	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.01)
Left	0.26	-	-
	(0.34)		
Right	-0.06	-	-
	(0.19)		
Parties in government	0.16*	0.16**	0.13*
	(0.09)	(0.08)	(0.07)
Treated	-0.64	-1.51	-1.49
	(0.79)	(0.92)	(0.91)
*2015law*Treated*	-1.30**	-1.07**	-0.94**
	(0.60)	(0.49)	(0.43)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Testing-down restriction	-	$\chi^2(19)=14.85$	$\chi^2(19)=14.85$
Observations	1,006	1,006	1,233
Number of municipalities	120	120	137

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. 'Larger sample' refers to the adding of extra observations when variables with missing observations do not survive the testing-down procedure. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B2: DiD estimation interacted with gender variable (general model, reduced model, and reduced model based on larger sample)

Dependent variable: *Social services spending* (in %)						
	% women in government			Female mayor		
	General	Reduced		General	Reduced	
		Smaller sample	Larger sample		Smaller sample	Larger sample
	(1)	(2)	(3)	(4)	(5)	(6)
Lag social services	0.70***	0.71***	0.68***	0.71***	0.72***	0.69***
	(0.11)	(0.11)	(0.13)	(0.12)	(0.13)	(0.13)
Own revenues	0.00	-	-	0.01	-	-
	(0.01)			(0.01)		
Lag own revenues	0.02	-	-	0.02	-	-
	(0.02)			(0.02)		
Current transfers	0.00	-	-	-0.00	-	-
	(0.02)			(0.02)		
Lag current transfers	0.03	-	-	0.03	-	-
	(0.04)			(0.04)		
Capital transfers	-0.01	-	-	-0.01	-	-
	(0.02)			(0.02)		
Lag capital transfers	0.04*	-	-	0.04*	-	-
	(0.02)			(0.02)		
Lag debt	0.00	-	-	0.00	-	-
	(0.00)			(0.00)		
Lag rent	-0.00	-	-	-0.00	-	-
	(0.00)			(0.00)		
Lag unemployment rate of youth	0.12	-	-	0.11	-	-
	(0.12)			(0.11)		
Lag unemployment rate of adults	-0.01	-	-	-0.00	-	-
	(0.03)			(0.03)		
Lag unemployment rate of elderly	-0.00	-	-	-0.00	-	-
	(0.04)			(0.04)		

Lag population density	-0.00	-	-	-0.00	-	-
	(0.00)			(0.00)		
Lag Share of youth	-0.07	-	-	-0.08	-	-
	(0.19)			(0.17)		
Lag Share of retired	0.15	-	-	0.15	-	-
	(0.09)			(0.09)		
Lag Share of immigration	-0.15*	-0.20**	-0.18**	-0.17*	-0.21**	-0.20**
	(0.08)	(0.09)	(0.08)	(0.09)	(0.11)	(0.09)
Mayor's age	0.01	-	-	0.01	-	-
	(0.01)			(0.01)		
Mayor's secondary education	-0.13	-	-	-0.10	-	-
	(0.34)			(0.30)		
Mayor's higher education	-0.39	-	-	-0.36	-	-
	(0.32)			(0.30)		
Female mayor	0.08	0.07	-0.22	0.80	0.62	0.12
	(0.46)	(0.47)	(0.52)	(0.50)	(0.58)	(0.66)
Female government	0.01	0.01	-0.00	-0.01	-0.01	-0.01
	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)
Left	0.19	-	-	0.26	-	-
	(0.34)			(0.37)		
Right	-0.02	-	-	-0.08	-	-
	(0.23)			(0.26)		
Parties in government	0.22**	0.22**	0.17**	0.14	-	-
	(0.11)	(0.10)	(0.08)	(0.09)		
2015-law	1.31	1.06	0.97	0.02	-	-
	(1.41)	(1.23)	(1.20)	(0.54)		
Treated	1.68	0.65	-0.22	-0.67	-1.51*	-1.50
	(2.45)	(2.35)	(1.99)	(0.73)	(0.92)	(0.94)
*2015law*Treated*	-6.65*	-6.29*	-5.02*	-1.30**	-0.82**	-0.77**
	(3.73)	(3.45)	(2.61)	(0.56)	(0.36)	(0.37)
*2015law*Female government*	-0.04	-0.03	-0.03	-	-	-
	(0.03)	(0.03)	(0.02)			

*Treated*Female government*	-0.05	-0.04	-0.03	-	-	-
	(0.04)	(0.04)	(0.03)			
*2015law*Treated*Female government*	0.13*	0.12*	0.10*	-	-	-
	(0.08)	(0.07)	(0.05)			
*2015law*Female mayor*	-	-	-	-1.40**	-1.16**	-0.96*
				(0.63)	(0.56)	(0.50)
*Treated*Female mayor*	-	-	-	0.52	0.55	0.96
				(1.34)	(1.41)	(1.47)
*2015law*Treated*Female mayor*	-	-	-	-0.22	-0.55	-0.53
				(2.04)	(1.92)	(1.76)
Time-trend	-0.06	-0.08	-0.04	-	-	-
	(0.13)	(0.10)	(0.09)			
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	Yes	Yes	Yes
Year trend	Yes	Yes	Yes	No	No	No
Testing-down restriction	-	$\chi^2(19)=15.15$	$\chi^2(19)=15.15$	-	$\chi^2(20)=21.82$	$\chi^2(20)=21.82$
Observations	1,006	1,006	1,233	1,006	1,006	1,233
Number of municipalities	120	120	137	120	120	137

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. 'Larger sample' refers to the adding of extra observations when variables with missing observations do not survive the testing-down procedure. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. [†] Time FE cannot be used to compute marginal effects and we use instead a linear trend together with variable *2015law* to control for time.

Table B3: DiD estimation on gender-equality and male-dominated government sample (general model, reduced model, and reduced model based on larger sample)

Dependent variable: *Social services spending* (in %)							
	Gender-equality					Male-dominated	
	Baseline			Female mayor		Baseline	
	General	Reduced		General	Reduced	General	Reduced
		Smaller sample	Larger sample		Smaller sample		Smaller sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Lag social services	0.15	0.22	0.10	0.15	0.21	0.75***	0.77***
	(0.23)	(0.25)	(0.13)	(0.23)	(0.25)	(0.13)	(0.14)
Own revenues	0.02	-	-	0.03	0.03*	-0.01	-
	(0.02)			(0.02)	(0.02)	(0.02)	
Lag own revenues	-0.01	-	-	-0.00	-	0.05	-
	(0.01)			(0.01)		(0.05)	
Current transfers	0.03	-	-	0.03	-	-0.02	-
	(0.02)			(0.02)		(0.04)	
Lag current transfers	-0.01	-	-	-0.01	-	0.07	-
	(0.01)			(0.01)		(0.07)	
Capital transfers	-0.01	-	-	-0.01	-	-0.00	-
	(0.02)			(0.02)		(0.03)	
Lag capital transfers	-0.00	-	-	0.00	-	0.05	-
	(0.01)			(0.01)		(0.04)	
Lag debt	0.00	-	-	0.00	-	-0.00	-
	(0.00)			(0.00)		(0.01)	
Lag rent	0.00	-	-	0.00	-	-0.00	-
	(0.00)			(0.00)		(0.00)	
Lag unemployment rate of youth	-0.11	-	-	-0.10	-	0.19	-
	(0.14)			(0.14)		(0.17)	
Lag unemployment rate of adults	0.06***	0.03*	0.04*	0.05**	0.03*	-0.03	-
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)	
Lag unemployment rate of elderly	-0.04	-	-	-0.02	-	-0.04	-

	(0.03)			(0.04)		(0.07)	
Lag population density	-0.01	-	-	-0.01	-	0.01	-
	(0.01)			(0.01)		(0.01)	
Lag Share of youth	-0.01	-	-	-0.05	-	-0.05	-
	(0.09)			(0.10)		(0.25)	
Lag Share of retired	-0.02	-	-	0.04	-	0.52***	0.61**
	(0.11)			(0.13)		(0.19)	(0.31)
Lag Share of immigration	-0.15**	-0.16**	-0.23	-0.14**	-0.17**	-0.18	-
	(0.06)	(0.07)	(0.18)	(0.06)	(0.07)	(0.19)	
Mayor's age	-0.03	-	-	-0.03	-	0.05*	0.07**
	(0.03)			(0.03)		(0.02)	(0.03)
Mayor's secondary education	0.48	-	-	0.59	-	-0.40	-
	(0.94)			(0.90)		(0.62)	
Mayor's higher education	0.67	-	-	0.67	0.38	-0.83	-
	(0.66)			(0.66)	(0.28)	(0.69)	
Female mayor	0.88	0.82	-0.17	2.35**	2.04**	-0.84	-1.17
	(0.85)	(0.72)	(0.65)	(1.15)	(0.91)	(0.98)	(0.86)
Left	-0.10	-	-	0.20	-	0.36	-
	(0.43)			(0.45)		(0.76)	
Right	-0.65	-	-	-0.51	-	0.54	-
	(0.48)			(0.48)		(0.67)	
Parties in government	-0.07	-	-	-0.09	-	0.79***	0.86***
	(0.10)			(0.10)		(0.22)	(0.23)
Treated	-1.12	-2.16	-4.00	-0.91	-2.15*	-	-
	(0.99)	(1.42)	(2.75)	(1.00)	(1.22)		
*2015law*Treated*	-0.19	-0.23	-1.06	-0.81	-0.77	-6.74**	-6.33*
	(0.42)	(0.42)	(0.80)	(0.60)	(0.49)	(3.20)	(3.33)
*2015law*Female mayor*	-	-	-	-1.24*	-1.19*	-	-
				(0.67)	(0.61)		

*Treated*Female mayor*	-	-	-	-2.95**	-2.64***	-	-
				(1.21)	(0.95)		
*2015law*Treated*Female mayor*	-	-	-	1.54*	1.88**	-	-
				(0.87)	(0.85)		
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Testing-down restriction	-	$\chi^2(19)=25.81$	$\chi^2(19)=25.81$	-	$\chi^2(17)=22.10$	-	$\chi^2(18)=16.29$
Observations	452	452	560	452	452	460	460
Number of municipalities	71	71	80	71	71	70	70

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. 'Larger sample' refers to the adding of extra observations when variables with missing observations do not survive the testing-down procedure. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table B4: DiD estimation conditional on the electoral term (general model, reduced model, and reduced model based on larger sample)

Dependent variable: *Social services spending* (in %)								
	All municipalities			Gender-equality			Male- dominated	
	General	Reduced		General	Reduced		General	Reduced
		Smaller sample	Larger sample		Smaller sample	Larger sample		Smaller sample
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag social services	0.72***	0.73***	0.70***	0.14	0.22	0.10	0.75***	0.77***
	(0.12)	(0.13)	(0.13)	(0.23)	(0.25)	(0.13)	(0.13)	(0.14)
Own revenues	0.00	-	-	0.02	-	-	-0.01	-
	(0.01)			(0.02)			(0.02)	
Lag own revenues	0.02	-	-	-0.00	-	-	0.05	-
	(0.02)			(0.01)			(0.05)	
Current transfers	-0.00	-	-	0.03	-	-	-0.02	-
	(0.02)			(0.02)			(0.04)	
Lag current transfers	0.03	-	-	-0.01	-	-	0.08	-
	(0.04)			(0.01)			(0.07)	
Capital transfers	-0.01	-	-	-0.01	-	-	-0.00	-
	(0.02)			(0.02)			(0.03)	
Lag capital transfers	0.03*	-	-	0.00	-	-	0.06	-
	(0.02)			(0.01)			(0.05)	
Lag debt	0.00	-	-	0.00	-	-	-0.00	-
	(0.00)			(0.00)			(0.01)	
Lag rent	-0.00	-	-	0.00	-	-	-0.00	-
	(0.00)			(0.00)			(0.00)	
Lag unemployment rate of youth	0.11	-	-	-0.10	-	-	0.19	-
	(0.11)			(0.15)			(0.17)	
Lag unemployment rate of adults	0.00	-	-	0.06***	0.03*	0.04*	-0.03	-
	(0.03)			(0.02)	(0.02)	(0.02)	(0.03)	

Lag unemployment rate of elderly	-0.02	-	-	-0.04	-	-	-0.04	-
	(0.04)			(0.03)			(0.07)	
Lag population density	-0.00	-	-	-0.01	-	-	0.01	-
	(0.00)			(0.01)			(0.01)	
Lag Share of youth	-0.05	-	-	-0.01	-	-	-0.05	-
	(0.17)			(0.09)			(0.25)	
Lag Share of retired	0.15	-	-	-0.02	-	-	0.53***	0.62**
	(0.09)			(0.12)			(0.19)	(0.31)
Lag Share of immigration	-0.16*	-0.21**	-0.20**	-0.15**	-0.16**	-0.23	-0.18	-
	(0.09)	(0.10)	(0.09)	(0.06)	(0.07)	(0.18)	(0.19)	
Mayor's age	0.01	-	-	-0.03	-	-	0.05*	0.07***
	(0.01)			(0.03)			(0.02)	(0.03)
Mayor's secondary education	-0.01	-	-	0.51	-	-	-0.47	-
	(0.29)			(0.91)			(0.60)	
Mayor's higher education	-0.26	-	-	0.71	-	-	-0.91	-
	(0.28)			(0.65)			(0.64)	
Female mayor	0.22	0.14	-0.15	0.89	0.84	-0.13	-0.73	-1.13
	(0.50)	(0.50)	(0.53)	(0.83)	(0.70)	(0.63)	(0.93)	(0.86)
Female government	-0.01	-0.01	-0.01	-	-	-	-	-
	(0.01)	(0.01)	(0.01)					
Left	0.26	-	-	-0.08	-	-	0.23	-
	(0.34)			(0.43)			(0.73)	
Right	-0.05	-	-	-0.63	-	-	0.55	-
	(0.20)			(0.48)			(0.67)	
Parties in government	0.16*	0.16**	0.13*	-0.08	-	-	0.79***	0.87***
	(0.09)	(0.08)	(0.07)	(0.10)			(0.22)	(0.24)
Pre-election	0.11	0.04	-0.15	-0.20	-0.13	0.01	-0.01	-0.07
	(0.34)	(0.30)	(0.33)	(0.18)	(0.18)	(0.17)	(0.72)	(0.65)
Treated	-0.54	-1.45	-1.32	-1.04	-2.05	-3.72	-	-
	(0.80)	(0.94)	(0.93)	(1.01)	(1.45)	(2.78)		
*2015law*Treated*	-1.53**	-1.20**	-1.20**	-0.37	-0.44	-1.42	-7.99**	-7.24**
	(0.68)	(0.53)	(0.54)	(0.48)	(0.42)	(0.93)	(3.35)	(3.50)
*2015law*Pre-election*	-	-	-	-	-	-	-	-

*Treated*Pre-election*	-0.50	-0.20	-0.66	-0.31	-0.43	-1.10	-3.44***	-2.58**
	(0.50)	(0.41)	(0.63)	(0.60)	(0.52)	(0.71)	(1.06)	(1.04)
*2015law*Treated*Pre-election*	0.90	0.61	1.10*	0.78	0.91	1.32*	4.16***	3.31***
	(0.57)	(0.50)	(0.60)	(0.59)	(0.60)	(0.69)	(1.23)	(1.23)
Municipality FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Testing-down restriction	-	$\chi2(19)=15.26$	$\chi2(19)=15.26$	-	$\chi2(19)=26.19$	$\chi2(19)=26.19$	-	$\chi2(18)=19.29$
Observations	1,006	1,006	1,233	452	452	560	460	460
Number of municipalities	120	120	137	71	71	80	70	70

Notes: Values are based on the dynamic Breitung et al. estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. 'Larger sample' refers to the adding of extra observations when variables with missing observations do not survive the testing-down procedure. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Appendix C: Classic DiD Fixed Effects estimator—Tables

Table C1: DiD estimation interacted with gender variable

Dependent variable: *Social services spending* (in %)		
	% women in government†	Female mayor
	(1)	(2)
*2015law*Treated*	-7.90**	-2.39**
	(3.76)	(0.99)
*2015law*Treated*Female government*	0.16*	-
	(0.08)	
*2015law*Treated*Female mayor*	-	0.52
		(2.96)
Municipality FE	Yes	Yes
Time FE	No	Yes
Year trend	Yes	No
Controls	No	No
Observations	1,370	1,370
Number of municipalities	137	137

Notes: Values are based on standard Fixed Effects estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1. † Time FE cannot be used to compute marginal effects and, to control for time, we use a linear trend together with variable *2015law*

Table C2: DiD estimation on gender-balanced and male-dominated government sample

Dependent variable: *Social services spending* (in %)			
	Gender-balanced		Male-dominated
	Baseline	Female mayor	Baseline
	(1)	(2)	(3)
*2015law*Treated*	-1.21	-0.95	-3.99*
	(0.79)	(0.66)	(2.08)
*2015law*Treated*Female mayor*	-	1.79*	-
		(0.97)	
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	No	No	No
Observations	689	534	531
Number of municipalities	104	100	103

Notes: Values are based on standard Fixed Effects estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

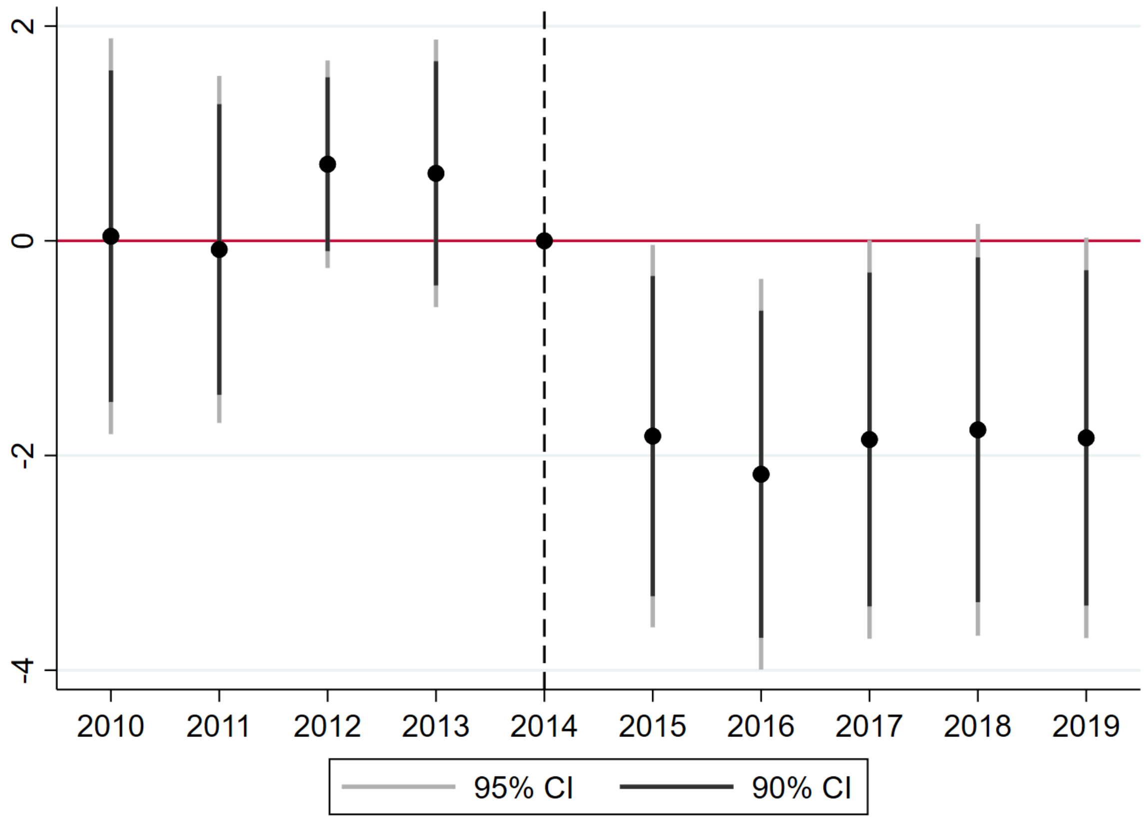
Table C3: DiD estimation conditional on the electoral term

Dependent variable: *Social services spending* (in %)			
	Full sample	Gender-balanced	Male-dominated
	(1)	(2)	(3)
*2015law*Treated*	-2.35**	-1.59*	-4.06*
	(0.91)	(0.91)	(2.06)
*2015law*Treated*Pre-election*	0.57*	1.05*	0.43
	(0.30)	(0.59)	(0.90)
Municipality FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	No	No	No
Observations	1,370	689	521
Number of municipalities	137	104	103

Notes: Values are based on standard Fixed Effects estimator for the sample of municipalities in which the elected mayor was of the same sex during the 2011 and 2015 elections. Municipality cluster-robust standard errors are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

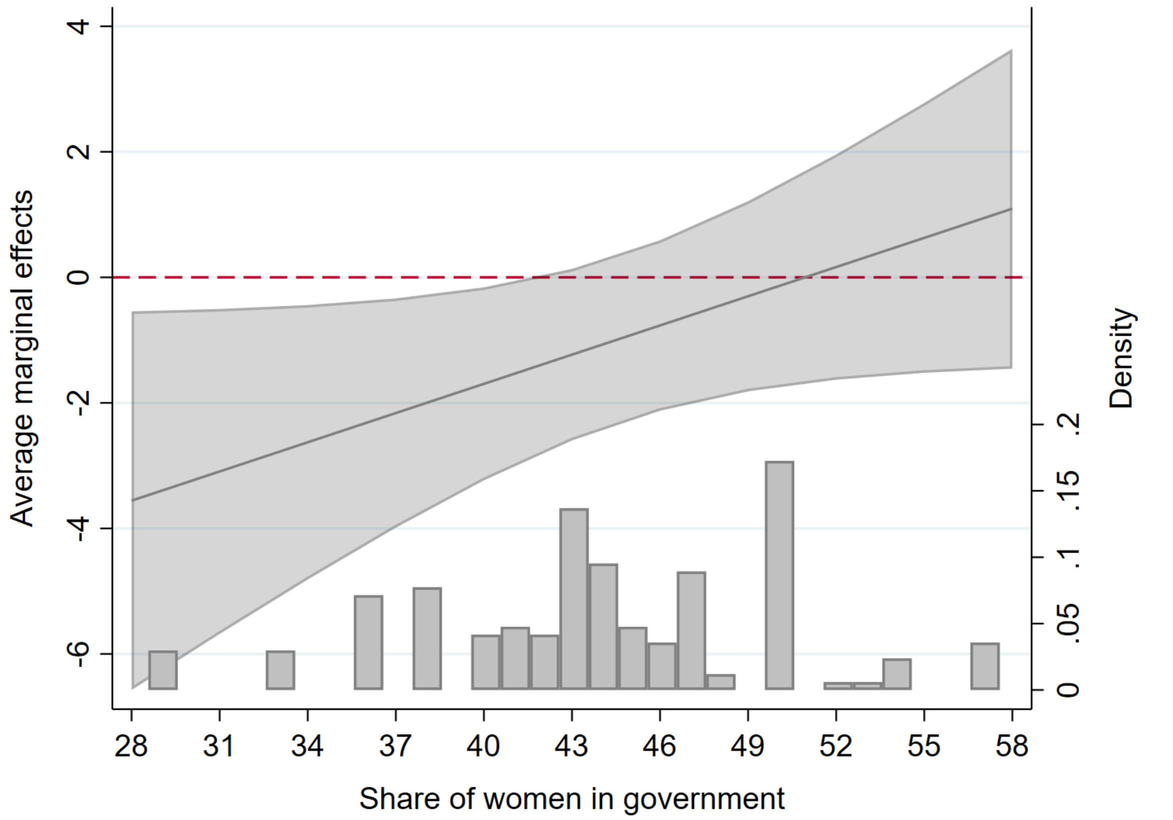
Appendix D: Classic DiD Fixed Effects estimator—Figures

Figure D1: Baseline event-study DiD estimation



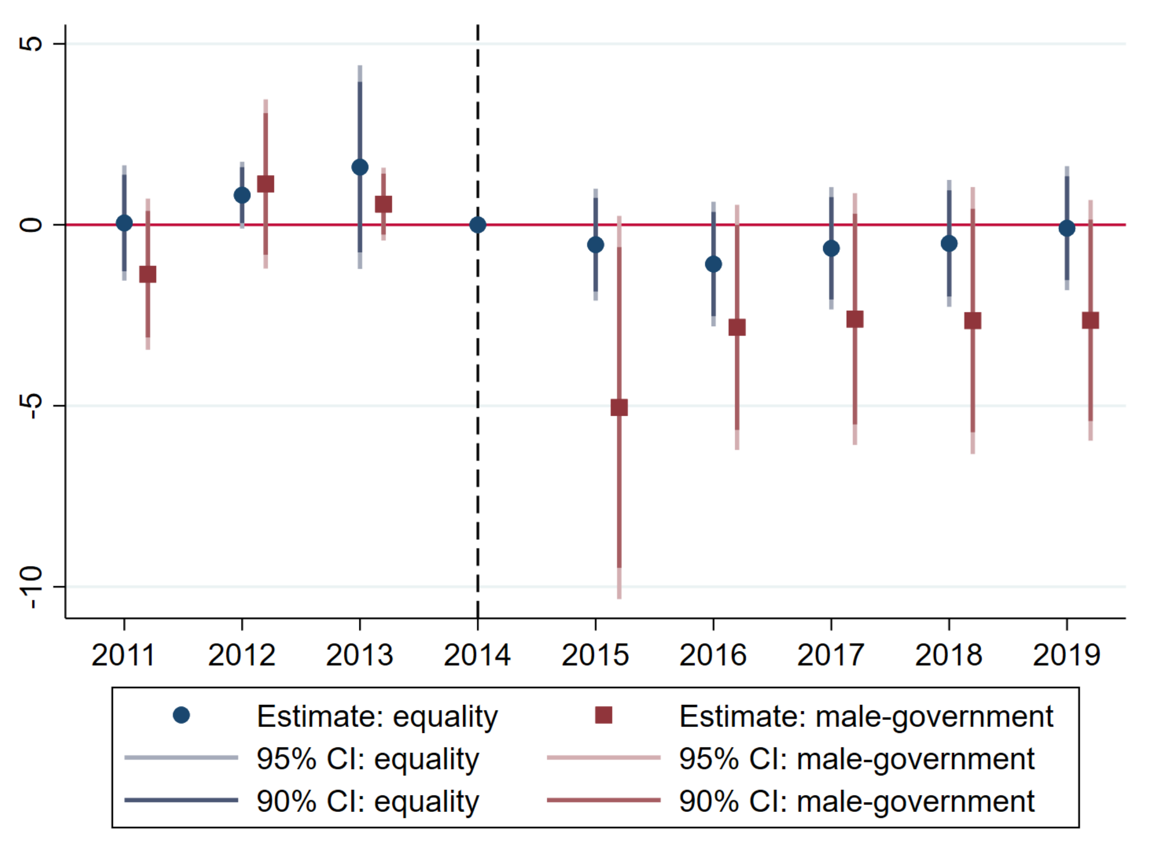
Notes: We estimate the baseline model reported in column 1 of Table 1, but *2015law* is substituted by year dummies. The effect is set to 0 in the last year before the reform, and the resulting reference year 2014 is marked by a dashed vertical line. Each dot in the graph shows the estimated effect, and the bars indicate 95% and 90%, respectively.

Figure D2: AMEs of treated municipalities conditional on the share of females politicians in government



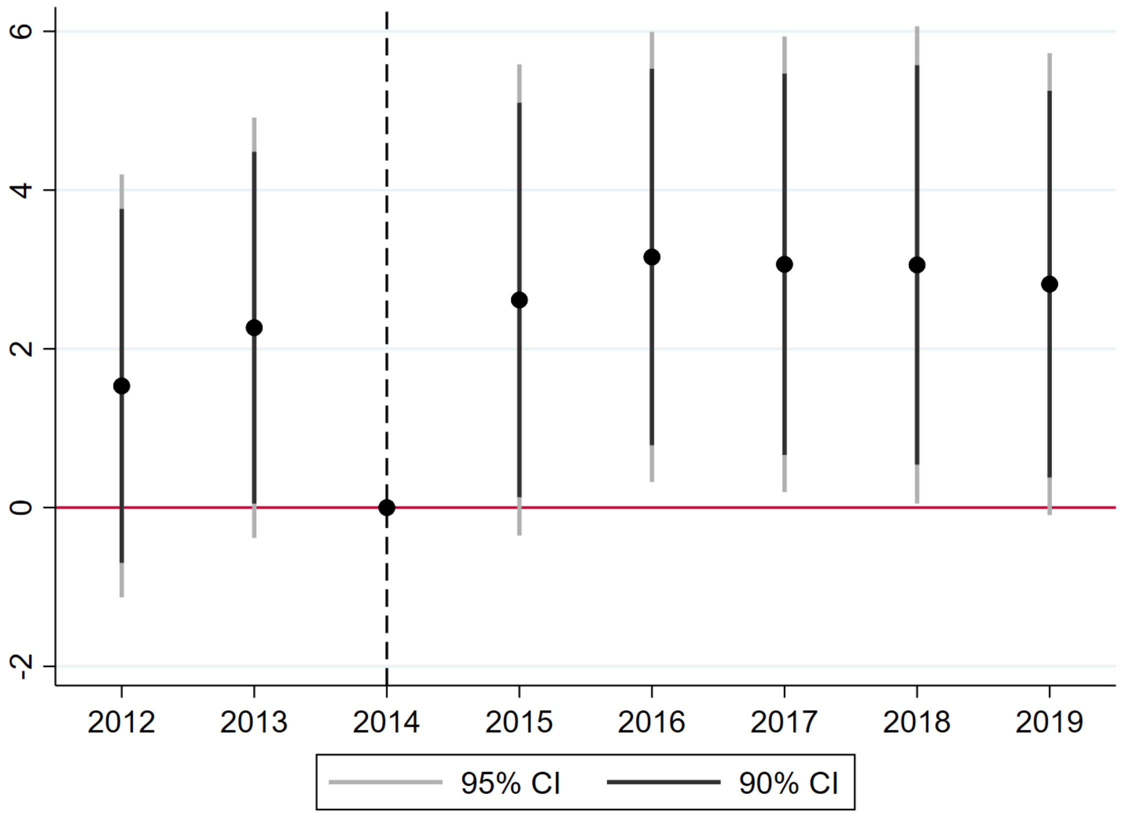
Notes: The point estimates with 95% confidence intervals show the AMEs for treated municipalities after the reform at different values of *Female government*, ranging from 28% (lowest) to 58% (highest). The graphs' bars indicate the density of the variable, *Female government* and are measured on the right-hand y-axis. The underlying regression is reported in Table C1, column 1.

Figure D3: Baseline event-study DiD estimation in both subsamples



Notes: We estimate the baseline models reported in Table C2, columns 1 and 3, 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 (and square) in the graph shows the estimated effect, and the bars indicate 95% and 90% confidence intervals, respectively.

Figure D4: Event-study DiD estimation with female mayor interaction



Notes: We estimate the baseline model interacted with *Female mayor* reported in Table C2, column 2, but variable *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 (and square) in the graph shows the estimated effect, and the bars indicate 95% and 90%, respectively.