# WOMEN'S REPRESENTATION IN POLITICS: VOTER BIAS, PARTY BIAS, AND ELECTORAL SYSTEMS

2018

Martín Gonzalez-Eiras and Carlos Sanz

Documentos de Trabajo
N.º 1834

BANCO DE ESPAÑA
Eurosistema

**WOMEN'S REPRESENTATION IN POLITICS: VOTER BIAS, PARTY BIAS, AND ELECTORAL SYSTEMS**

# WOMEN'S REPRESENTATION IN POLITICS: VOTER BIAS, PARTY BIAS, AND ELECTORAL SYSTEMS(*)

Martín Gonzalez-Eiras (**)

UNIVERSITY OF COPENHAGEN

Carlos Sanz (***)

BANCO DE ESPAÑA

(*) We thank Guillermo Caruana, Thomas Fujiwara, Nezih Guner, Mónica Martínez-Bravo, Enrique Moral-Benito, Roberto Ramos, Johanna Rickne, Roco Titiunik, Ernesto Villanueva, and seminar participants at the Banco de España, CEMFI, CUNEF, ESEM, IBEO workshop, Universidad Carlos III, and Universidad de Montevideo for helpful comments. We thank Andrés Gago and Carlos Victoria for sharing the data on the gender of mayors, Manuel Bagues for sharing the codes to match first names with gender, and Alfonso Gordaliza for invaluable information on Spanish municipal politics.
(**) Corresponding author; mge@alumn.mit.edu; Oster Farimagsgade 5, 1353 Copenhagen K, Denmark.
(***) carlossanz@bde.es; Alcalá 48 28014 Madrid, Spain.

Documentos de Trabajo. N.º 1834

2018

The Working Paper Series seeks to disseminate original research in economics and finance. All papers have been anonymously refereed. By publishing these papers, the Banco de España aims to contribute to economic analysis and, in particular, to knowledge of the Spanish economy and its international environment.

The opinions and analyses in the Working Paper Series are the responsibility of the authors and, therefore, do not necessarily coincide with those of the Banco de España or the Eurosystem.

The Banco de España disseminates its main reports and most of its publications via the Internet at the following website: http://www.bde.es.



ISSN: 1579-8666 (on line)

**Abstract**

We study how electoral systems affect the presence of women in politics using a model in which both voters and parties might have a gender bias. We apply the model to Spanish municipal elections, in which national law mandates that municipalities follow one of two different electoral systems: a closed-list system in which voters pick one party-list, or an open-list system, in which voters pick individual candidates. Using a regression discontinuity design, we find that the closed-list system increases the share of women among candidates and councilors by 2.5 percentage points, and the share of women among mayors by 4.3 percentage points. Our model explains these results as mostly driven by voter bias against women. We provide evidence that supports the mechanism of the model. In particular, we show that, when two councilors almost tied in general-election votes, the one with "one more vote" is substantially more likely to be appointed mayor, but this does not happen when the most voted was female and the second was male, suggesting the presence of some voter bias. We also show that, in a subsample of municipalities with low bias — proxied by having had a female mayor in the past — the difference between the two electoral systems disappears.

**Keywords:** voting, electoral systems, gender bias, regression discontinuity.

**JEL classification:** D72, J16, J71.

**Resumen**

Estudiamos cómo afecta el sistema electoral a la presencia de mujeres en política usando un modelo en el que tanto los votantes como los partidos pueden tener un sesgo de género. Aplicamos el modelo a las elecciones municipales españolas, en las que, según la ley, los municipios utilizan dos sistemas electorales distintos: uno de listas cerradas, en el que los votantes votan por un partido, o uno de listas abiertas, en el que votan por candidatos individuales. Usando un análisis de regresión discontinua, encontramos que el sistema de listas cerradas aumenta 2,5 puntos porcentuales la proporción de mujeres entre los candidatos y los concejales, y 4,3 puntos porcentuales entre los alcaldes. Según el modelo, estos resultados se explican por un sesgo de los votantes. Proporcionamos evidencia que apoya el mecanismo del modelo. En particular, mostramos que, con carácter general, cuando dos concejales casi empatan en votos, es considerablemente más probable que el que obtuvo «un voto más» sea nombrado alcalde, pero este resultado no se observa con generalidad cuando el más votado fue una mujer y el segundo un hombre, lo que sugiere la presencia de un cierto sesgo de género. También mostramos que, en una submuestra de municipios con poco sesgo —aquellos que han tenido una alcaldesa en el pasado—, la diferencia entre los dos sistemas electorales desaparece.

**Palabras clave:** votación, sistemas electorales, sesgos de género, regresión discontinua.

**Códigos JEL:** D72, J16, J71.

# 1 Introduction

Women are underrepresented in politics. For example, they hold only 23% of legislative seats in national assemblies.[1] This situation has attracted much attention both in the media and in the academic literature. Perhaps the most important reason behind this interest is that having women in public office may have an impact on policy. There is evidence, for example, that female leaders cater more to women's needs.[2] In addition, women in public office may serve as role models, improving women's self-confidence and affecting the attitudes of citizens towards women inside and outside politics.[3]

How to increase the presence of women in politics is therefore a pressing question. In recent decades, several countries have introduced gender quotas that require parties to field a minimum of female candidates. There is consensus that quotas are successful in increasing female participation in politics, but they have some drawbacks.[4] In this paper we show, both theoretically and empirically, that electoral systems may have a sizable impact on women's representation in the presence of voter or party biases against women.[5]

Our analysis is based on Spanish municipal elections. We exploit the fact that, depending on their population size, municipalities follow one of two different electoral systems. Municipalities with more than 250 inhabitants must use a closed list (CL), proportional representation system to elect a city council. Municipalities with 250 or fewer inhabitants must use an open list (OL) system, in which voters can vote for up to four individual candidates from the same or different parties.[6] In the first meeting after the election, the council selects a mayor among the councilors. In the CL system, only the councilors at the top of their party-lists can be selected, while in the OL system any councilor can be selected. Therefore, these two systems differ starkly: in CL, voters pick party-lists, while in OL, they pick individual candidates. Municipalities at both sides of the population threshold are unaffected by gender quotas.[7]

This setup provides several attractive features. First, the electoral system is determined by population size, as mandated by *national* law. This allows us to introduce a regression discontinuity design, which is well known for providing credible estimates of causal effects (Lee and Lemieux 2010), and compare women's representation in the two electoral systems. Second, the treatment is clearly defined, as all the municipalities in one or the other electoral system must follow the same rules. This is in opposition to cross-country studies, where it is inevitable to pool different systems into the same category. Third, no other regulation changes at the threshold. Therefore, the differences between municipalities at each side of the threshold can be confidently attributed to the electoral system and not to some other fact. Finally, there are many observations close to the threshold. There are around 2,400 municipalities in a window of 150 inhabitants, and data are available for four election-years (2003, 2007, 2011, and 2015).

[1] World Bank data as of 2017. See *https : //data.worldbank.org/indicator/SG.GEN.PARL.ZS*.

[2] For example, Chattopadhyay and Duflo (2004) show that female leaders spend more in infrastructure that is directly relevant to the needs of women.

[3] See, for example, Beaman, Chattopadhyay, Duflo, Pande, and Topalova (2009) and Beaman, Duflo, Pande, and Topalova (2012).

[4] See the discussion in Pande and Ford (2011). First, quotas can reduce women's incentives to invest, if they believe that their advancement is made easier by quotas. Second, quotas can be perceived as unfair and worsen attitudes towards women—women selected through quotas may be stereotyped as less qualified. Third, they can even fail to lift the barriers that prevent women from playing an influential role in politics. Bagues and Campa (2017a) study, also in Spain, how the introduction of quotas in larger municipalities affected women's presence in politics. They find that quotas increased the share of female candidates and councilors, but failed to increase the share of women among party leaders or mayors.

[5] By the electoral system, we mean the set of rules that determines how elections are conducted and their results determined, especially how ballots are marked, cast, and counted.

[6] More specifically, given that voters can vote for candidates from different party-lists, it is a *free list* system. Throughout the paper, we stick to the more general term *open list*.

[7] The 2007 Equality Law introduced gender quotas for Spanish elections, but only for municipalities of more than 3,000 inhabitants, which are far from the threshold and are therefore not used in our analysis.

We develop a model for candidate choice and mayor selection in which there are two types of agents—voters and parties—which might have a gender bias.[8] There is first a general election to elect the councilors. After this election, elected councilors select among themselves a mayor, who makes policy choices. Parties are office motivated and have to choose the candidates for the general election from a limited supply of potential candidates, who differ in gender and competence. We will say that there is *voter bias* if voters have a preference against women that makes them prefer lower-competence men to more competent women. The size of the bias captures the extent to which they are willing to trade off competence and gender. We will say that there is *party bias* if parties have a preference against women that makes them prefer to increase the probability that a man is in office at the expense of reducing the probability of winning the election. The size of the bias captures the extent to which they are willing to trade off the probability of winning and gender.

Hence, parties may decide to choose fewer female candidates because they have a preference against women (party bias), or because they internalize voter bias, knowing that biased voters are less likely to vote for female candidates. Given that the mayor makes policy choices, we assume that both voters and parties care about the competence and quality *of the mayor*. Thus, in CL, voters and parties may have a bias against the top-listed candidate, who will become mayor if the party wins the election. In contrast, in OL, voters may have a bias against all candidates in the list, since voters do not know ex ante who will be later selected as mayor.

We show, both analytically and with numerical simulations, that the effect of the electoral system on the share of female candidates, councilors, and mayors depends on the underlying source of gender bias. In particular, in the presence of voter bias and no party bias, the CL system results in a higher share of female councilors, while, in the presence of party and no voter bias, the opposite is the case. In the CL system, both voter and party bias will lead parties to choose fewer female candidates for the top spot of the list (as only the top-listed candidate can become mayor). In the OL system, however, the incentives are different under voter and party bias. With voter bias, parties have an incentive to choose fewer female candidates for all positions in the list, as biased voters will penalize them. Hence, in this case, OL will result in fewer female councilors than CL. With party bias, parties have an incentive to choose one male candidate (so that he may be appointed mayor later) but will complete the list with the most competent candidates regardless of their gender. Thus, the OL system results in a higher share of female councilors than CL, as voters are unbiased in their choice among selected candidates.

We test the model predictions in the Spanish context. Regression discontinuity (RD) estimates show that, relative to the OL system, the CL system increases the share of female candidates by between 2 and 3 percentage points (p.p.), a relative increase of 8-12%. Similarly, the CL system increases the share of female councilors by between 2 and 3 p.p. (9-14%) and the share of female mayors by between 3.5 and 6 p.p. (22-37%). According to the model, these results imply that voter bias is more relevant than party bias, as the model predicts that, in the presence of voter bias, the CL system increases the share of female candidates, councilors, and mayors. By contrast, as explained above, in the presence of party bias, the CL system should *decrease* the share of female councilors.

We present three sets of results that support the mechanism of the model. First, using experience as a control for competence, female candidates attract fewer votes than male candidates, suggesting the presence of voter bias against women.[9] Having a female candidate for mayor in a CL election reduces the vote share of that list by 3.3 p.p., relative to having a male candidate. In OL, female candidates attract 1.3 p.p. less votes than male candidates. We also show that the gender and quality of the top-listed candidate in the CL system are more relevant than those of lower-ranked candidates, as assumed in the model.

---

[8]By bias we do not necessarily mean taste-based discrimination (Becker 1957), but rather whatever reasons why voters or parties may have a net preference for men when all the other observable variables are kept constant.

[9]These results hold after including municipality and party-year fixed effects, which may account for many unobservable factors that could affect the composition of the lists and the votes attracted by that list. However, given that experience is admittedly a crude measure of competence, these regressions cannot be totally conclusive on the presence of voter bias.

Second, we study the mayor selection in the OL system, and provide evidence that (i) there is a social norm that the councilor with most general election votes should become mayor: even when the two most voted councilors almost tied in votes, the councilor with "one more vote" is 19 p.p. more likely to be appointed mayor; (ii) there are gender differences in the norm: it is followed when the top two councilors are of the same gender, or when the first is male and the second is female, but not when the first is female and the second is male, and (iii) voters enforce the norm and its gender differences by punishing, in the next election, parties that deviate from it, further suggesting the presence of voter bias against women.

Third, we test an additional prediction of the model—that, in the absence of voter and party bias, the two electoral systems should lead to the same share of female candidates, councilors, or mayors. To do so, we focus on a subsample of municipalities—those that have been exposed to a female mayor in the past—in which we conjecture that voter or party biases are weaker. We first show that, unlike what happens in the whole sample, female candidates do *not* attract fewer votes than male candidates in the subsample of exposed municipalities, suggesting that, as conjectured, voter and party biases are smaller in this subsample.[10] We then estimate our RD regressions in this subsample and find that the positive effect of CL on the share of female candidates, councilors, and mayors disappears, as predicted by the model.

While the evidence is consistent with the mechanism proposed in the model, we discuss other possible explanations to our findings, and conclude that they are unlikely to drive the results. First, we analyze whether the effects can be driven by the difference in district magnitude—there are seven councilors in CL, and five in OL. To assess whether district magnitude has an effect separate from the effect of the ballot structure, we exploit the fact that, at 1,000 inhabitants, there is another threshold, at which there is a change in district magnitude but not in ballot structure. We perform a similar RD analysis and find no significant effects, against the hypothesis that the district magnitude drives the results. Second, we discuss why the supply of candidates (women's willingness to run) cannot explain the entirety of our results. Third, we show that our findings are unlikely to be driven by different party structures in the two electoral systems.

Our paper contributes to three strands of the literature. First, we add to the literature on the effects of electoral systems on women's representation in politics. Previous work has found that CL systems increase women's representation relative to OL systems (e.g., Thames and Williams (2010) and Valdini (2012)).[11] We add to this literature both empirically and theoretically. On the empirical side, we contribute by providing evidence from a setting that allows us to obtain a clean identification. Estimating the effects of electoral systems is challenging due to endogeneity issues.[12] For this reason, exploiting discontinuities at population thresholds is becoming a popular way to obtain credible estimates of causal effects. To the best of our knowledge, our paper provides the first RD to estimate the effect of electoral systems on women's representation. On the theoretical side, the literature has discussed the importance of voter and party biases and how they can interact with the electoral system, but we were still lacking a formal model that explicitly incorporates both voter and party biases. We contribute by developing such model. Although we lay out the specifics

[10]We do not claim causality in that exposure to a female mayor reduces bias. It is indeed possible that exposure to a female leader lowers voter bias (Beaman, Chattopadhyay, Duflo, Pande, and Topalova 2009), and it can reduce party bias if elected mayors influence party politics at higher levels. But it is also possible that exposed municipalities had a lower bias initially.

[11]Thames and Williams (2010) uses OLS cross-country regressions and Valdini (2012) exploits the reform of the electoral system in Japan in 1994 to compare women's representation before and after the reform. In contrast, Schmidt (2009), using OLS cross-country regressions, does not find significant effects.

[12]Roberts, Seawright, and Cyr (2013) argue that most previous literature on the effects of electoral systems on women's representation suffers from serious methodological problems. They try to circumvent the issue using within-country changes in electoral systems and matching methods, and find that the effect of electoral systems may not be as large as previously thought. Similarly, Fortin-Rittberger and Rittberger (2014) argue that the empirical connection between electoral systems and descriptive representation might in fact be an endogenous rather than a causal relationship.

to match the Spanish context and obtain predictions for our empirical analysis, the model is flexible and can be applied to other settings.[13]

Second, our paper contributes to the literature that studies the reasons behind women's underrepresentation in politics. The literature has discussed three main possibilities: (i) that women's underrepresentation is supply-driven, with women shying away from politics, (ii) that it follows from voter bias against women, and (iii) that it follows from party bias against women. We provide evidence that allows us to disentangle voter and party biases. The results indicate that, in our context, voter bias plays a more important role. These findings are in line with Fréchette, Maniquet, and Morelli (2008) and De Paola, Scoppa, and Lombardo (2010), who suggest that the lack of female legislators may reflect voter bias in France and Italy, respectively. However, they differ from the findings in Esteve-Volart and Bagues (2012), who find that party bias is more prevalent in elections to the Spanish Senate.[14]

Finally, we add to the growing literature on social norms—informal understandings that govern individual behavior in society—in political environments. On this front, our paper is most closely related Fujiwara and Sanz (2018).[15] They provide evidence of a norm that "the most voted *party* should form the government in a parliamentary system", and show that such norm has important implications for policy and the design of electoral systems. In this paper, we show that a similar norm exists in bargaining *among councilors of the same party*, namely, that "the most voted *councilor* should become mayor". We go further by showing that the norm applies differently by gender: the norm is followed when the two most voted councilors are of the same gender or when the first is male and the second is female, but not when the first is female and the second is male. We also provide evidence that, similarly to the norm in Fujiwara and Sanz (2018), the norm we identify is enforced by voters, who punish in the next election parties that do not follow it. Voters also enforce the gender differences: the punishment is larger (smaller) when the first is male (female) and the second is female (male). In this regard, our findings are related to the growing literature that suggests that men and women are punished differently for the same behavior or outcome.[16]

The rest of this paper is organized as follows. Section 2 provides some background on the institutional setting. Section 3 lays out the theoretical model and numerical simulation results. Section 4 presents the data, the empirical strategy, and the results on the effect of the electoral system on women's representation. Section 5 presents the additional results that support the mechanism of the model. Section 6 discusses other possible mechanisms. Section 7 concludes. The appendix contains proofs, details of the numerical simulations, and ancillary discussions

## 2 Background: Spanish Municipal Elections

Spain is a highly decentralized country. It is divided into 17 regions and more than 8,000 municipalities. Each municipality is run by a municipal government that has substantial autonomy:

[13]As examples, it can be adapted to compare women's representation in majoritarian and proportional representation systems, or to study the impact of double-preference voting conditioned on gender, as recently introduced in Italian local elections.

[14]There are two possible reasons behind this difference. First, we study municipal elections while Esteve-Volart and Bagues (2012) study national elections, and biases can operate differently, even within the same electorate. In particular, voters may have been more exposed to women at national politics than at local politics, reducing their bias in national elections. This explanation is consistent with Dahl, Kotsadam, and Rooth (2018), who find that exposure to rank-and-file women in the military in Norway does not have a spillover effect on attitudes about higher up female leadership, i.e. what is required is exposure to female leaders at the relevant level. Second, the municipalities that we study are small and therefore carry little weight in the regressions of Esteve-Volart and Bagues (2012). If bias against women is more prevalent in small municipalities, then this can explain the different findings.

[15]Other related papers on the importance of norms are Anagol and Fujiwara (2016), who show that second-placed candidates are substantially more likely than close third-placed candidates to run in, and win, subsequent elections, and Folke, Persson, and Rickne (2016), who show that narrow "list winners" of preferential voting in a general election in Sweden are over 50% more likely to become party leaders than their runner-ups.

[16]For example, female financial advisors are more likely to be fired for misconduct than male advisors (Egan, Matvos, and Seru (2017)), and female surgeons are punished harder (receive less referrals) than male surgeons following a patient death (Sarsons (2017)).

they can set their own taxes (the most important being a property tax) and spend a considerable amount of money (municipalities close to the threshold spend more than 600 euros per capita per year on average—see Sanz (2017a)).[17] It is important to note that, even though the 2007 Equality Law mandates gender quotas, this is of no direct consequence for our study as only municipalities of more than 3,000 inhabitants are subject to the law.[18]

Municipal elections are held simultaneously in all municipalities every four years. The electoral system depends on the population size of the municipality in the year before municipal elections: municipalities with a population of more than 250 inhabitants must use a CL proportional representation, and those with 250 or fewer (but at least 100) inhabitants must use an OL system.[19] This institutional framework was established by national law in 1978, before the first municipal elections after Franco's regime, and has not changed since then.[20]

Municipalities in the CL system elect a city council in a single-district election. All municipalities in this system that are used for identification elect a seven-member council (larger municipalities have larger councils, but, again, this is irrelevant for our purposes). Each party presents a list of candidates and voters pick one of the party-lists.[21] To convert votes into seats the D'Hondt rule is used.[22] Councilors are drawn from each list in the order in which the candidates are listed. For example, if a party obtains four seats according to the votes obtained by the party-list, then the four candidates at the top of the list become councilors. In the first meeting after the election, the council elects a mayor among the councilors. We call this election the "mayor selection". In CL, only councilors that were at the first spot of their party-lists can be appointed mayor. This means that parties choose their candidate for mayor before the general election.[23] Although the council is responsible for approving the budget, controlling the governing bodies, and for the roll-call vote of confidence on the mayor, mayors are the "*the center of gravity of political life in the municipality*" and by law hold the most important executive functions and exercise leadership in municipal politics (Vallés and Brugué 2001).[24] Mayors have a central role in running the government by chairing council meetings and appointing and dismissing cabinet members and staff, and have substantial control over determination and allocation of expenditures, since they prepare municipal budgets and approve construction processes. Indeed, Spanish municipal governments exemplify a case of strong executive power and have been described as municipal presidentialism (Magre-Ferran and Bertrana-Horta 2005).

Municipalities in the OL system elect a city council of five members. Candidates are listed on party-lists created by political parties, as in the CL system. However, voters do not pick one of the party-lists, but rather check up to four candidates belonging to one or more party-lists. The five most voted individual candidates are elected members of the council.[25] As in the CL system, there is a "mayor selection" in the first council meeting after the general election, in which the council

[17]Municipal governments spend 13% of the overall spending of the country. Typical services that are provided by small municipalities include public lighting, waste collection, cemeteries, street cleaning, road pavement, touristic information to visitors, and local festivities.

[18]The threshold was 5,000 in 2007. The effect of this law is studied by Bagues and Campa (2017a), Bagues and Campa (2017b), and Casas-Arce and Saiz (2015).

[19]Municipalities with fewer than 100 inhabitants use a completely different *government* system, namely, a direct democracy system—see Sanz (2017a).

[20]The OL system for smaller municipalities was introduced after a proposal by the right-wing Alianza Popular, and was only opposed by the Communist Party. The argument to introduce the OL system only in small municipalities was that it is more likely that people know the candidates in those municipalities than in larger places.

[21]Typically, it is the provincial management of political parties that is responsible of making the lists. Provincial managers have delegates at the municipal level, who talk to potential candidates and hence provide the provincial managers with first-hand input to help create the lists.

[22]There is an electoral threshold at 5%, i.e. parties need to get at least 5% of the votes to enter the D'Hondt distribution of seats. However, this threshold very rarely plays a role in the municipalities of interest here, due to the low number of seats.

[23]Throughout the paper, we use the expressions "candidate for mayor" and "top-listed candidate" interchangeably.

[24]Mayors are very rarely removed by the council, further indicating their strength. In our sample, 96.5% of mayors stayed the whole term.

[25]Technically, this is a *limited voting* system, as candidates can vote for a maximum number of candidates that is lower than the number of seats in the council.

selects a mayor among its members. Importantly, unlike in CL, all of the councilors can be elected as mayors, not only those at the top spot of the party-lists. The roles of the council and the mayor are identical under the two systems.

# 3 Theoretical Model

## 3.1 Setup

We consider a static setting with three types of agents: voters, candidates, and political parties. There are two parties, $A$ and $B$, that differ in their ideology and compete in an election. Voters vote in a general election to elect the councilors. Each party presents a list of two candidates, and three councilors are elected.[26] We consider two electoral systems: CL and OL. Under CL, voters choose the entire list of candidates from one party. Under OL, voters cast two preference votes which are not restricted to be for the same party. Parties have to choose the two candidates in their lists. Candidates are in limited supply and each party can choose from four options, two male and two female. We assume that the competence, or quality, of candidates is i.i.d., drawn from the same distribution and that, while it is perfectly observed by parties, it is only imperfectly observed by voters.

After voting takes place, there is a mayor selection, in which the three elected councilors elect a mayor among themselves. The mayor will implement policy during the term.[27] Note that, given that there are three elected councilors and only two parties, one party will always have the majority in the council. This matches the reality of elections in small Spanish municipalities.[28] Only councilors at the top of the party-lists can be elected mayors in CL. Hence, the mayor will always be the councilor at the top of the list of the party that has the majority in the council, i.e. the party that obtained the most votes in the general election. In OL, by contrast, any councilor can be elected mayor. Hence, the mayor will be either of the two councilors of the majority party.

There is a continuum of voters that care about the quality and gender of candidates and the ideology of parties.[29] We assume preferences are aggregated through a standard probabilistic voting model as in Lindbeck and Weibull (1987). An explicit microfoundation is provided in appendix A. To illustrate the results, we consider the comparison of two candidates, one from each party, $A1$ and $B1$. We assume that the indirect utilities that voter $i$ receives from these candidates are given by

$$v^i(A1) = s^{A1} - \mu \mathbb{1}_{g(A1)=f} + \sigma^i + \delta, \quad (1)$$

$$v^i(B1) = s^{B1} - \mu \mathbb{1}_{g(B1)=f}, \quad (2)$$

where $s^x$ is the competence of candidate $x$, and $g(x)$ is its gender, $f$ for female and $m$ for male. The parameter $\mu \geq 0$ measures the average gender bias in the electorate.[30] Since voters observe candidate competence imperfectly, $\sigma^i$ captures both the uncertainty in candidates competence, and the importance of ideology from voting for party $A$. It is assumed that $\sigma^i$ is uniformly distributed with mean zero.[31] Voters' preferences are also affected by a common relative popularity shock, $\delta$,

[26]In Spanish municipal elections, as discussed, seven city councilors are elected in CL and five in OL. We argue in the discussion of alternative mechanisms that this difference in the number of seats does not drive the results.

[27]In Spain, the mayor is much more influential than the other members of the council, as discussed in Section 2.

[28]In our sample, in municipalities under the OL system one party has a majority in 97.5% of the councils. In municipalities under CL, this happens in 92.2% of the councils.

[29]Under citizen-candidate models, such as Osborne and Slivinski (1996) and Besley and Coate (1997), voters care about candidate quality as this will result in the implementation of better policies. One interpretation is that the probability that a candidate solves a problem that benefits voters is increasing in her competence.

[30]For simplicity, we assume that every voter has the same bias. We could instead have different biases for male and female voters, or a distribution of biases. As long as the bias, or its distribution, is orthogonal to the distribution of ideology, results are not affected.

[31]Since we have only one group of voters, this distribution, which renders the probability of winning a voter's support a continuous function of the candidate's characteristics, becomes irrelevant. Thus, we do not need to specify it.

which is also assumed to be uniformly distributed, $\delta \sim U[-\frac{1}{\psi}, \frac{1}{\psi}]$, with $\psi > 0$.

With these assumptions, and considering a vote restricted to take place only between $A1$ and $B1$, the probability that party $A$ wins is given by (see appendix A for derivation):

$$P^A = \frac{1}{2} + \psi \left( s^{A1} - s^{B1} - \mu(\mathbb{1}_{g(A1)=f} - \mathbb{1}_{g(B1)=f}) \right). \tag{3}$$

Given that the mayor will implement policy and that only top-listed candidates can be elected mayors in CL, voters will only care about the quality and gender of top-listed candidates in CL.[32] Under this assumption, equation (3) gives the probability that party $A$ wins the election under the CL system.

In OL, we assume that voters first rank candidates in each party such that the candidate ranked first for party $i$ is the one with highest value for competence net of gender bias. Then a pairwise comparison of the highest ranked candidates for both parties, $A1$ and $B1$, is made, and the first vote is cast for the candidate that gives the voter the higher utility according to (1) and (2). To cast the second vote, the voter now compares the second ranked candidate of the party that got her first vote with the highest ranked candidate of the other party, i.e. a comparison of either $A2$ and $B1$, or $B2$ and $A1$. Again the vote is cast to the candidate that gives the voter the higher utility according to (1) and (2). The three most voted candidates are elected and the party that gets both of its candidates elected chooses the mayor among them. Importantly, since voters only imperfectly observe candidate competence, and thus do not know ex ante which candidate will be chosen as mayor, they will express a gender bias when casting *both* votes in the OL system. This stands in contrast to the CL system, where only the gender of the top-listed candidate matters, as only this candidate can eventually be appointed mayor.

We assume that parties are office motivated and might also have a bias against women. In CL, given that only top-listed candidates can be elected mayors, we assume that party $j$ maximizes

$$U^{jCL} = P^j - \nu\psi \mathbb{1}_{g(j1)=f},$$

where $\nu\psi \geq 0$ is the cost to parties, relative to the rents from office, of choosing a female candidate for mayor.

In OL, by contrast, any councilor can be elected mayor. Thus, we have to characterize the mayor selection process. We follow Fujiwara and Sanz (2018) in assuming that elections serve the purpose of aggregating information. Thus, election outcomes in the OL system inform voters on the quality of the candidates. As a result, voters have a preference that the most voted candidate of the party with two councilors (as opposed to the second most voted candidate of that party) is appointed mayor. But parties are not constrained in this choice, creating a potential conflict of interest between them and voters. The latter thus have a reason to punish, *in the next election*, a party that does not select their currently preferred candidate. Thus, we assume that the objective function of parties in OL has an extra term capturing the future expected consequences of the current mayor choice:

$$U^{jOL} = P^j - \nu\psi \mathbb{1}_{g(ma(j))=f} - \beta\psi \mathbb{1}_{ma(j)=j2} \left[ 1 - \mu\alpha \mathbb{1}_{g(j1)=f} \right], \tag{4}$$

where $ma(j)$ denotes party $j$'s choice of mayor, $\mathbb{1}_{ma(j)=j2}$ is an indicator function that takes the value of 1 when the party chooses as mayor the second most voted candidate, and $\beta\psi \geq 0$ is the cost, in terms of future election prospects relative to current rents in office, of not choosing the most voted candidate as mayor in the OL system.[33] Note that we allow for gender differences in

[32]In section 5.1, we show evidence that voters care significantly more about the competence and gender of the top-listed candidate relative to those of other candidates. Here, for tractability, we assume that they care *only* about the quality and gender of the top-listed candidate. For a formal analysis, see appendix A.

[33]For simplicity, we assume that imperfect information on candidates' competence does not affect their relative ranking in the vote tally, i.e. $j1$ will get more votes than $j2$.

the norm, i.e. that the norm may be enforced differently depending on whether the most voted candidate is a male or female. For the former, the relative cost is $\beta\psi$, while for the latter, it is $\beta\psi(1-\mu\alpha)$. Thus, voter bias reduces the cost of not selecting the most voted candidate if this happens to be female.

Hence, the model assumes that, if both candidates are of the same gender, then the most voted one will be selected as mayor to avoid punishment in future elections. The same happens if the most voted candidate is male and the second most voted candidate is female. But if the most voted candidate is female and the second most voted candidate is male, then the party will select the male candidate as mayor if and only if

$$\beta\left[1-\mu\alpha\right]<\nu.$$

### 3.2 Analytical Results

We start by evaluating the predictions of the model regarding the electoral impact of having female (as opposed to male) candidates. For the CL system, we study the impact of having a woman, as opposed to a man, as the candidate for mayor, i.e. as the top-listed candidate. Equation (3) gives the probability that a party wins an election, conditional on the competence and gender of the competing candidates. Therefore, if party $A$ is contemplating replacing $m_1$ with $f_1$ as its candidate for mayor, this has an effect on its vote share that is proportional to:

$$P^A|f_1-P^A|m_1=\psi\left(s^{f1}-s^{m1}-\mu\right).$$

The effect of replacing a male by a female candidate on the electoral prospects in OL is derived similarly using (15), the equivalent of (3). We find that if party $A$ is contemplating replacing, say, $m_1$ with $f_2$ as a candidate, this has an effect on its vote share proportional to:

$$P^A|f_1,f_2-P^A|f_1,m_1=\frac{\psi}{2}\left(s^{f2}-s^{m1}-\mu\right).$$

We summarize these result in the following proposition.

**Proposition 1.** *If there is voter bias against women ($\mu>0$), then having a female candidate for mayor in CL (female candidates in OL) reduces the party's vote share, conditional on candidates' competence.*

To evaluate the effect of the electoral system on the share of female candidates, councilors, and mayors, we specify how the supply of candidates is determined. Candidates' competence is assumed to be distributed according to $H(\cdot)$.[34] Denote by $m_1$ and $m_2<m_1$ (respectively $f_1$ and $f_2<f_1$) the two competence draws for male (female) candidates from this distribution for a given party. We denote by $H_1$ ($H_2$) the probability distribution for the maximum (minimum) of two independent draws. The probability that the party chooses a female candidate for mayor under the CL system is given by

$$\begin{aligned}P(f_1>m_1+\mu+\nu) &= \int_0^\infty\int_{m_1+\mu+\nu}^\infty dH_1(f_1)dH_1(m_1)\\ &= \int_0^\infty\left(1-H_1(m_1+\mu+\nu)\right)dH_1(m_1)\\ &= 1-\int_0^\infty H_1(m_1+\mu+\nu)dH_1(m_1).\end{aligned}$$

Since $H_1(\cdot)$ is an increasing function, the probability of choosing a female candidate is decreasing with voter and party bias against women.

[34] We only impose minor restrictions on this distribution function, namely, that it has a compact support and no mass points. Furthermore, these properties should be inherited by the distributions of the maximum and minimum of two independent draws.

In the CL system, we assume that the party will choose the most competent candidate, among the remaining three draws, for the second place in the list.[35] Denoting by $F$ the number of female candidates, the probabilities that the party chooses none, $P^{CL}(F=0)$, or two, $P^{CL}(F=2)$, female candidates are given by[36]

$$\begin{aligned} P^{CL}(F=0) &= P(f_1 < m_1 + \mu + \nu \wedge f_1 < m_2) = P(f_1 < m_2), & (5) \\ P^{CL}(F=2) &= P(f_1 > m_1 + \mu + \nu \wedge f_2 > m_1) = P(m_1 < \min[f_2, f_1 - \mu - \nu]), & (6) \end{aligned}$$

and the probability of one female candidate is $P^{CL}(F=1) = 1 - P^{CL}(F=0) - P^{CL}(F=2)$. Equations (5) and (6) show that, in CL, parties may have a bias against having a female candidate for mayor, and also internalize voter bias when comparing $f_1$ and $m_1$, but there is no bias for the second placed candidate.

In the OL system, parties choose candidates taking into consideration voter gender bias for both candidates and that the *ex post* mayor selection might entail a cost if the most voted candidate is female and the party selects the male second most voted candidate. Thus,[37]

$$\begin{aligned} P^{OL}(F=0) &= P(f_1 < m_2 + \mu + 2\mathbb{1}_{f_1 > m_1 + \mu} \min[\nu, \beta(1-\mu\alpha)]) \\ &= P(f_1 < m_2 + \mu) + P(m_1 + \mu < f_1 < m_2 + \mu + 2\min[\nu, \beta(1-\mu\alpha)]), & (7) \\ P^{OL}(F=2) &= P(f_2 > m_1 + \mu + 2\nu - 2\min[\nu, \beta(1-\mu\alpha)]), & (8) \end{aligned}$$

where $\mathbb{1}_{f_1 > m_1 + \mu}$ is an indicator for whether the highest ranked female candidate would get more votes than the highest ranked male candidate. Parties, when comparing the most competent candidates, know that choosing one female and one male candidates gives them the option of later selecting the male if he finishes second.[38] The probability of one female candidate, $P^{OL}(F=1)$, is given by $1 - P^{OL}(F=0) - P^{OL}(F=2)$. The difference in the expected number of female candidates across electoral systems is given by

$$\begin{aligned} \mathrm{E}[F]^{CL} - E[F]^{OL} &= P^{CL}(F=1) + 2P^{CL}(F=2) - \left[P^{OL}(F=1) + 2P^{OL}(F=2)\right], \\ &= \left[P^{OL}(F=0) - P^{CL}(F=0)\right] + \left[P^{CL}(F=2) - P^{OL}(F=2)\right]. \end{aligned}$$

**Proposition 2.** *If there is voter bias against women ($\mu > 0$) and no party bias ($\nu = 0$), then the CL system increases the expected share of female candidates ($\mathrm{E}[F]^{CL} - E[F]^{OL} > 0$). If there is party bias ($\nu > 0$), no voter bias ($\mu = 0$), $\nu < \beta$, and $\nu << 1$, then the expected share of female candidates is approximately the same under both electoral systems ($\mathrm{E}[F]^{CL} \approx E[F]^{OL}$). If there is party bias ($\nu > 0$), no voter bias ($\mu = 0$), and $\nu >> E[f_1] - E[f_2]$, then the CL system increases the expected share of female candidates ($\mathrm{E}[F]^{CL} - E[F]^{OL} > 0$).*

*Proof.* See appendix B.

To determine the effect of the electoral system on the share of female *councilors*, we need to combine the results of propositions 1 and 2, as a party's electoral performance determines whether it gets one or two councilors, and performance depends on candidates' competence and gender.

[35] Recall that, in CL, voters care only about competence and gender of top-listed candidates. We assume that parties follow this convention to reward competent candidates, and preserve them for future elections.

[36] The expressions below involve integrals of $dH_1$ and $dH_2$ as in the derivation of $P(f_1 > m_1 + \mu + \nu)$ above.

[37] The following expressions come from the comparison of objective function (4) for the three possible cases, choosing both male candidates, both female, or a male and a female. Taking into account the probability of wining the election (equation 15), these options have the following values for parties:

$$\begin{aligned} &m_1 + m_2, \\ &f_1 + f_2 - 2\nu - 2\mu, \\ &f_1 + m_1 - \mu - 2\mathbb{1}_{f_1 > m_1 + \mu} \min[\nu, \beta(1-\alpha\mu)], \end{aligned}$$

where the indicator $\mathbb{1}_{f_1 > m_1 + \mu}$ shows that when the female candidate will get fewer votes than the male candidate, the party can chose the male candidate as mayor with no cost.

The case of voter bias will be explored with numerical simulations. For the case of party bias, but no voter bias, candidates' gender has no effect on vote shares. Proposition 2 shows that, for low values of party bias, such that $\nu < \beta$, the share of female candidates is the same across electoral systems. This does not imply that the share of female councilors is the same in CL and OL, because in the former there will be less female top-listed candidates due to party bias. Since only top-listed candidates become councilors for the losing party, this implies that the share of female councilors is lower in CL.

**Proposition 3.** *If there is party bias against women ($\nu > 0$), no voter bias ($\mu = 0$), $\nu < \beta$, and $\nu << 1$, then the OL system increases the expected share of female councilors.*

Finally, we study the model's implications for the share of female *mayors*. We start by recalling that the probability that a party chooses a female candidate for mayor in CL is given by

$$P^{CL}(g(ma(j)) = f) = P(f_1 > m_1 + \mu + \nu). \tag{9}$$

The analysis is again complicated since the probability of having a female mayor also depends on the probability that a party that chooses a female candidate for mayor wins the election, $E[P^j | f_1 > m_1 + \mu + \nu]$.

**Proposition 4.** *If there is voter bias against women ($\mu > 0$), no party bias ($\nu = 0$), and $\beta(1-\alpha\mu) > \nu = 0$, then the expected share of female mayors is approximately the same under the two systems. If there is party bias ($\nu > 0$), no voter bias ($\mu = 0$), and $\nu < \beta$, then the expected share of female mayors is higher under the OL system. If there is party bias ($\nu > 0$), no voter bias ($\mu = 0$), and $\nu > \beta$, then the expected share of female mayors is higher under the CL system.*

*Proof.* See appendix B.

### 3.3 Numerical Simulations

To complete the analytical predictions provided so far, we resort to numerical simulations (for details see appendix C). In the simulations, we make $\beta$ and $\alpha$ stochastic such that, in equilibrium, there is variation in the degree with which the norm is followed when the most voted candidate is female. Figure 1 shows the results of these simulations, both for voter bias (panel A), and party bias (panel B). Three different strengths for the norm are considered, "strong" when the probability that a woman is not appointed mayor after winning the popular vote is low, "intermediate" when this probability is approximately 15%, and "weak" when it is around 30%. Results confirm the predictions of propositions 2, 3, and 4.

In panel A, it can be seen that in the presence of voter bias and no party bias, the CL system leads to a higher share of female candidates and councilors than the OL system, and the gap is increasing in voter bias. These effects are very similar in the three graphs of Panel A, indicating that they do not depend on the strength of the norm. The CL system also increases the share of female mayors. This effect is, as expected, very sensitive to the strength of the norm.

Panel B of figure 1 considers the case of party bias when there is no voter bias. For low values of party bias, the two systems result in approximately the same share of female candidates, and, as party bias increases, CL results in relatively more female candidates. In contrast, for low values of party bias, the CL system leads to *fewer* female councilors than the OL system. Finally, the effect on the share of female mayors depends on the strength of the norm. If the norm is strong, then the OL system increases the share of female mayors. If the norm is weak or intermediate, CL results in a higher share of female mayors than OL.

# 4 Empirical Evidence

In this section, we empirically estimate the effect of the electoral system on the share of female candidates, councilors, and mayors. Section 4.1 lays out the data used in the regressions, section 4.2 presents the empirical strategy, and section 4.3 shows the results.

## 4.1 Data

Data for municipal elections are from the Ministry of the Interior and are publicly available. Information on the gender of candidates, councilors, and mayors can be obtained for four election-years: 2003, 2007, 2011, and 2015. For municipalities in the OL system and for those in the CL system in 2003, we imputed the gender from the first name. Names in Spain are strongly gender-oriented, so ambiguous cases are extremely rare (this approach has been used by previous work, e.g. Casas-Arce and Saiz (2015) and Esteve-Volart and Bagues (2012)). Additionally, we use data from national Congress elections (also from the Ministry of Interior) and some demographic variables (from the National Institute of Statistics) to assess covariate balance around the threshold. These data are also publicly available.

We focus on municipalities close to the threshold, which we define as those within a window of 150 inhabitants around the cutoff, that is, from 100 to 400 inhabitants.[39] The final data set contains observations for 9,465 municipal elections.[40]

Table 1 shows the share of female candidates, councilors, and mayors, by year (panel A) and by electoral system (panel B). On average, 24.9% of candidates, 21.7% of councilors, and 15.3% of mayors were female. Three things are worth noting. First, women are more present among candidates than among councilors and mayors. This is because women are ranked lower in the lists (in CL), or because they receive fewer votes than male candidates (in OL).[41] Second, women's representation has increased significantly over the sample period. For instance, the share of female candidates rose from 20.8% in 2003 to 27.6% in 2015. Third, women are more represented in the CL than in the OL system. This *prima facie* evidence might suggest that the CL system increases women's representation, but should not be interpreted causally at this stage, as it is a mere comparison of means.

## 4.2 Empirical Strategy

We implement a RD design to estimate the effect of the electoral system on the share of female candidates, councilors, and mayors:

$$Outcome_{mt} = \alpha + \chi CL_{mt} + f(Pop_{mt} - 250) + u_{mt}, \tag{10}$$

where $Outcome_{mt}$ is the outcome of interest (e.g., the share of female candidates), $CL_{mt}$ is a treatment dummy that takes the value of 1 if municipality $m$ used the CL system in election-year $t$, $Pop_{mt}$ is the assignment variable (population the year before the elections), and $u_{mt}$ is an error term. The parameter of interest is $\chi$.

The identification assumption is that municipalities at both sides of the threshold do not differ in characteristics, other than the electoral system, that may affect women's representation in politics. We assess the validity of the empirical approach with two tests that are standard in RD designs.

First, we show that there is no manipulation around the threshold. If having a population size just above or just below the threshold is as good as random, then the density of populations should

[39] It is natural to consider 150 as the window of analysis because there is a change in the government system at 100 inhabitants, as explained in Section 2.

[40] This is 97.4% of all the potential observations. The remaining are cases with missing or erroneous data in the original files, or in which we cannot identify the gender of one or more candidates or councilors.

[41] Note that the position a person within the list can also affect the votes that one gets in the OL system: voters tend to vote for candidates that are at the top of the list—see Esteve-Volart and Bagues (2012).

be continuous near the threshold. We implement the test introduced by Cattaneo, Jansson, and Ma (2017) and cannot reject the null of no manipulation (p-value = 0.39), as shown in figure 2.[42]

Second, we show that covariates are balanced around the threshold. We study whether municipalities at both sides of the threshold are similar in a number of socioeconomic and political characteristics: average age of the population, share of foreigners and EU foreigners, a measure of the ideology (the difference in votes shares in the last Congress election before election-year $t$ between the two main parties, the right-wing Popular Party (PP) and the left-wing Socialist Party (PSOE)), voter turnout, and unemployment rate. We also study whether the observations at both sides of the threshold are equally likely to come from any given year. As mentioned, women's representation has been increasing over our sample period. If our CL observations are from more recent years than those from OL, that could spuriously drive our results. Hence, if the empirical strategy is valid, we should find that the electoral system is not correlated with any of these variables. The results, displayed in table 2 and figure 3, show that all of the coefficients are close to zero and not statistically significant at any conventional level.

## 4.3 Results

Table 3 presents the main estimates of the impact of the electoral system on the share of female candidates, councilors, and mayors. Column (1) provides the estimated effect with a linear polynomial and the bandwidth proposed by Imbens and Kalyanaraman (2012). Column (2) shows the results when the bandwidth by Calonico, Cattaneo, and Titiunik (2014) is employed. Columns (3) and (4) use the entire sample with linear and quadratic specifications. Column (5) uses the local randomization strategy introduced by Cattaneo, Frandsen, and Titiunik (2015).[43] Results for a wide range of bandwidths are shown in figure A1.

Panel A of Table 3 shows the effect of the electoral system on candidates. The CL system increases the share of female candidates by between 2 and 3 p.p. (8-12%). This effect is significant at conventional levels. A graphical representation of the results can be seen in panel A of figure 4. Panel B of table 3 and panel B of figure 4 show the results for councilors. The CL system increases the share of female councilors by between 2 and 3 p.p. (9-14%). This effect is also statistically significant. Finally, Panel C of table 3 and panel C of figure 4 show the effect on mayors. The results indicate that the CL system increases the share of female mayors by between 3.5 and 6 p.p (22-37%). The estimates are noisier than those for candidates and councilors, but significant in most specifications. An inspection to panel C of figure A1 also reveals estimates within this range for a large set of bandwidths, suggesting that the CL system does increase the share of female mayors.

In sum, the evidence indicates that the CL system increases the share of female candidates, councilors, and mayors. According to the model, these results are likely driven by voter bias, as (moderate) party bias has the counterfactual implication that we should observe a higher share of female councilors in OL than in CL. According to the numerical simulations reported in figure 1, voter bias of between 0.03 and 0.05 can account for the observed difference in female candidates and councilors between the two electoral systems.[44] For these values of voter bias to also account for the observed differences in the share of female *mayors*, it must be the case that the norm for

[42] This is in contrast to Sanz (2017b). Note, however, that our paper focuses on the final years of the period, so the samples are different. In particular, there is evidence that manipulation of population figures almost disappeared after 2005, due to an improvement in the enforcement of population numbers by the Statistics Agency (Foremny and Solé-Ollé (2016)).

[43] To calculate the window, we use all of the covariates listed in table 2. We start from the possible smallest window (size 1) and study 10 10-inhabitant increments. The resulting window is 31 inhabitants. Note that, under local randomization, one derives the exact distribution of the statistic in the randomization of the treatment. This distribution is exact and need not be symmetric, and hence standard errors are not defined. For this reason, we report the p-values in this column.

[44] The average competence of a candidate is normalized to be one. For the Pareto distribution used, the average competence of two independent draws is roughly 1.61. Thus, voter bias of between 0.03 and 0.05 corresponds to voter bias of between 1.8% and 3.1% of the average competence of the highest-ranked female candidate.

mayor selection is not followed with probability between 15% and 30% when a female candidate wins the popular vote. In the next section, we will provide some additional results that support the presence of voter bias in our context and back the mechanism of the model.

# 5 Additional Results

In this section, we provide additional results that lend support to the mechanism of the model. In Section 5.1, we (i) show that female candidates attract fewer votes than male candidates, suggesting the presence of voter bias; (ii) identify a subsample of municipalities—those that have been exposed to a female mayor in the past—in which female and male candidates attract the same votes; and (iii) provide support for the assumption of the model that, in the CL system, the gender and quality of the top-listed candidate are more relevant than those of lower-ranked candidates. In Section 5.2, we study the mayor selection in the OL system, and provide evidence that (i) as posited in the model, there is a norm that the councilor with most general election votes should become mayor; (ii) there are gender differences in this norm; and (iii) voters enforce the norm and its gender differences, further suggesting the presence of voter bias against women. In Section 5.3, we test an additional prediction of the model, namely, that in the absence of voter and party bias, the two systems should result in the same share of female candidates, councilors, and mayors. We provide evidence that is consistent with this prediction, by showing that the effect of CL is not present in the subsample of municipalities that have been exposed to a female mayor.

## 5.1 The Gender of Candidates and Vote Shares

In CL, we estimate the effect of the gender of the top-listed candidate (i.e. the candidate for mayor) on the votes received by the party-list:

$$VSh_{pmt} = \gamma(pt) + \gamma(m) + \theta_{fem} FemCandidate_{1,pmt} + \theta_{exp} ExpCandidate_{1,pmt} + \epsilon_{pmt}, \quad (11)$$

where $VSh_{pmt}$ is the vote share of party $p$ in municipality $m$ and election-year $t$, $FemCandidate_1$ is a dummy that indicates whether party $p$'s top-listed candidate is female, and $ExpCandidate_{1,pmt}$ is a dummy that indicates whether the top-listed candidate has been a councilor in the past.[45]. To avoid incumbency effects, we restrict the sample to non-incumbent parties.[46] In OL, we estimate the same regression, but the unit of observation is an individual candidate (instead of a party), i.e. we regress the vote share of a candidate on that candidate's gender and experience and the fixed effects.

Table 4 shows the results. In column (1), which does not include the experience control, we can see that having a female candidate for mayor in CL reduces the party's vote share by 3.8 p.p (panel A). In OL, a female candidate is expected to obtain 2 p.p. less than a male counterpart (column (1) of panel B). Both effects are statistically significant at the 1% level. Column (2) adds the experience control. The coefficients are slightly reduced compared with the unconditional results shown in column (1), but remain negative and significant at the 1% level, both in CL (panel A) and OL (panel B). As expected, the results also show that experienced candidates command a larger share of votes. Hence, according to proposition 1, these results suggest the presence of voter bias.[47]

We then split the sample based on whether the municipality has had a female mayor at any point during the two previous terms.[48] We expect that biases against women are lower in municipalities that have been exposed to a female mayor, for two reasons: (i) the fact that a female was elected

[45] Other papers have used experience as a proxy for competence (Esteve-Volart and Bagues (2012) and Galasso and Nannicini (2017)). Another obvious option is education, but unfortunately we do not observe the education of candidates.

[46] The results of these tests are very similar if we consider the outcome in differences ($VSh_{pmt} - VSh_{pmt-4}$) instead of levels, or if we consider a (less demanding) specification with separate party and time fixed effects, i.e. $\gamma(p) + \gamma(t)$ instead of $\gamma(pt)$.

[47] Of course, they should be interpreted with caution, given that experience is a crude measure of competence.

mayor indicates that they probably were municipalities with initially low voter and party biases, and (ii) having been exposed to a female mayor can lower voter bias (Beaman, Chattopadhyay, Duflo, Pande, and Topalova (2009)).[49] Consistent with our hypothesis, the negative effect of women candidates on vote shares is larger in non-exposed municipalities (column (3) of table 4) and is not present in exposed municipalities (column (4)). This pattern occurs in both CL (panel A) and OL (panel B). We will make use of this subsample division in section 5.3, to test the model's prediction that there should be no effect of the electoral system in municipalities with no bias against women.

Finally, we provide evidence that backs the model assumption that, in CL, voters only care about the gender and experience of the top-listed candidate (or that they care substantially more than about the gender and experience of the rest of the candidates in the list). In column (5) of panel A, we add to equation (11) the variables $FemCand_{2-7,pmt}$ and $ExpCand_{2-7,pmt}$, the gender and experience of candidates ranked in positions 2 to 7 of the party-lists. If the model's assumption is correct, the effect of $FemCand_{1,pmt}$ and $ExpCand_{1,pmt}$ should be substantially larger than those of $FemCand_{2-7,pmt}$ and $ExpCand_{2-7,pmt}$. The results support this assumption. While having a woman as the top-listed candidate reduces the vote share of the list by 3 p.p., this effect is 0.8 p.p. for a candidate ranked lower in the list. For experience, the effect is also larger for the top-listed candidate than for other candidates, although the gap is smaller (7 versus 5.7 p.p.).

### 5.2 Mayor Selection in OL

We begin by providing evidence that, as assumed in our model, there is a norm that the councilor that obtained the most votes in the general election is substantially more likely to be appointed mayor than the runner-up, even when the two councilors almost tied in votes. Intuitively, we want to compare the two most voted councilors in the list of the party that appointed the mayor, and estimate whether the one with more votes is more likely to become the mayor. For example, consider an election in which one party has three councilors and another party has two—hence, the decision of whom to appoint mayor corresponds in practice to the councilors of the party with three councilors.[50] The general-election votes for the councilors of this party were 50, 49, and 35. We want to compare the probability that the councilors with 50 and 49 votes are appointed mayor. Given that obtaining one more vote is essentially random, the two councilors with 50 and 49 votes should be, on average, equal in any other characteristic (education, quality, gender, etc.). Hence, if the councilor with 50 votes is more likely to become mayor, this identifies a norm that "the most voted councilor has to become mayor".[51]

We use RD to estimate the effects, following the empirical strategy of Fujiwara and Sanz (2018). We define the running variable as follows:

$$x_{cmt} = \begin{cases} v_{mt} & \text{if } c \text{ is the most voted} \\ -v_{mt} & \text{if } c \text{ is second most voted,} \end{cases} \tag{12}$$

where $c$ refers to a councilor of the party of the mayor, and $v_{mt}$ is the difference in general-election votes between the first and second most voted councilors of the party that appointed the mayor in period $t$.

Therefore, for each election, we have two observations, one for the most voted councilor of the party of the mayor, and another for the second most voted councilor. If $x_{cmt} > 0$, then councilor $c$

[48]We focus on the last two terms (8 years) to avoid including exposition that is very distant in the past. The results are, in any case, very similar if we split the sample based on having been exposed to a female mayor at any point in the past.

[49]Having a female mayor can also reduce party bias, if the mayor is sufficiently powerful in the relevant party structure, i.e. in the provincial or regional power structure behind candidate choices.

[50]In 97.5% of the OL elections, one party has a majority in the council.

[51]There is a provision in the law that stipulates that, if no councilor obtains the majority of the votes in the mayor selection in the council, then the councilor with the most general-election votes is appointed mayor by default. We believe that this is unlikely to explain the effect—see the discussion in Fujiwara and Sanz (2018) as to why a similar provision in their context cannot explain outcomes. Furthermore, this provision of the law obviously does not differentiate by gender and hence cannot explain the gender differences that we identify.

has the most votes (*First*), and it has the second most votes otherwise. Let $y_{cmt}$ be the outcome, that is, becoming the mayor. The effect of having most votes is given by $\lim_{x_{cmt}\downarrow 0} \mathrm{E}[y_{cmt}|x_{cmt}] - \lim_{x_{cmt}\uparrow 0} \mathrm{E}[y_{cmt}|x_{cmt}]$, which can be estimated with a RD design:

$$y_{cmt} = \theta_0 + \theta_1 First + f(x_{cmt}) + \epsilon_{cmt}. \tag{13}$$

The identification assumption is that barely being the first or second most voted councilor does not correlate with any other factor that affects whether that councilor is appointed mayor. Intuitively, we are comparing two councilors of the same party, one of which obtained just one more vote than the other. As long as the number of votes that candidates are going to obtain cannot be precisely controlled, obtaining one more or less vote should be as good as random and the identification assumption is likely to hold. In particular, the two most voted councilors of the mayor's party should be equal in any characteristics.[52] As a validity check, we show in figure A2 that the first and second most voted councilors are in fact equally likely to be female. Also note that, by design, all covariates that do not vary within a municipality-year are balanced, and that the standard errors are not affected by the double-counting of elections, as they are clustered by municipality (see Fujiwara and Sanz (2018) for more details on the identification approach).[53]

The results from estimating equation (13) are in column (1) of table 5 and displayed graphically in figure 5.[54] When the two most voted councilors in the mayor's party-list almost tie in votes, the most voted councilor is 18.6 p.p. more likely to be appointed mayor than the second most voted councilor. The remaining columns of table 5 (and panels b-e of figure 5) study gender differences in the norm, by splitting the sample depending on the gender of the two most voted councilors. They reveal significant differences: the norm is strongest when the second most voted candidate is female, and weakest (non-existent) when the first is female and the second male. In appendix D, we provide an additional test of the gender differences in this norm.

We now turn to investigate the mechanism behind the norm (and its gender differences). We can think of two potential explanations. The first is that the norm arises from an agreement within parties to save bargaining costs. Every time a party gets a majority in the council, it needs to decide which of its councilors to appoint as mayor. Having the norm that "the most voted councilor is appointed mayor" can facilitate the decision process. If parties are biased against women, this could also account for the gender differences in the norm. The second potential explanation is that the norm is enforced by voters. Fujiwara and Sanz (2018) provide evidence that, in a variety of contexts, voters think that it is more "democratic" or "fair" that the most voted party forms the government. They also develop a model that shows that such norm can maximize welfare. As discussed in section 3, a norm that the most voted councilor is appointed mayor would also maximize welfare in our context, as the most voted councilor is expected to be of higher quality than the second most voted.

Although we cannot rule out that the first explanation is at play to some degree, we argue that the latter explanation is most likely. First, Fujiwara and Sanz (2018) provide compelling evidence that voters enforce a similar norm, also in Spain, by punishing in the next election parties that break the norm. It is reasonable to think that a similar norm in the same country is also enforced by voters. Second, and more importantly, we do a similar analysis and find that voters also punish

[52] One issue that can affect the votes obtained by candidates in the OL system is their position within the party-list, as there is evidence that voters tend to vote more for candidates that appear higher in the list (e.g. Esteve-Volart and Bagues (2012) show this happens in elections to the Spanish Senate). We cannot observe the positions of candidates within party-lists but this should also be balanced near the threshold and not drive the results. The logic is the same as for any predetermined characteristic. If obtaining one more or fewer votes is essentially random, the candidate that obtained one more votes than the other should be, on average, equally likely to be the candidate placed at the top of the list.

[53] To avoid incumbency effects, we restrict the sample to non-incumbent parties, i.e. we only consider mayor selection by those parties that were not in office in the previous period. The results are very similar if we do not make this restriction.

[54] These results are based on the IK bandwidth, but are robust to other bandwidth choices, as can be seen in figure A3.

parties that break the norm in our context. In particular, we test how the vote share of a party that follows the norm evolves in the next election compared with a party that breaks the norm:

$$VSh_{mt+4} - VSh_{mt} = \gamma(p) + \gamma(m) + \gamma(t) + \phi Norm_{mt} + \xi Norm * GenderDiff_{mt} + \epsilon_{mt}, \quad (14)$$

where $VSh_{pmt}$ is the vote share of the party that appointed the mayor in municipality $m$ in election-year $t$, $Norm_{mt}$ is a dummy that indicates whether the norm is followed in the mayor selection—that is, it takes the value of 1 if the most voted councilor of the party of the mayor is appointed mayor and 0 otherwise—and $\gamma(p)$, $\gamma(m)$, and $\gamma(t)$ are party, municipality and election-year fixed effects.[55] The sample is restricted to municipalities in which the party did not have the mayor at the previous term (i.e., time $t-4$), to avoid incumbency effects. The coefficient $\phi$ captures how following the norm in the mayor selection benefits the incumbent in the next election. Note that gaining votes from following the norm is equivalent to losing votes for not following the norm—we can only estimate the *difference* between following and not following the norm. If voters enforce the norm, we expect $\phi$ to be positive.

To study whether gender differences in the norm are also enforced by voters, we add the control *Norm∗GenderDiff*, where *GenderDiff* is a variable that takes the value of 1 if the most voted councilor was female and the second was male, 0 if they were of the same gender, and -1 if the most voted was male and the second was female. Hence, when this control is added, the coefficient on *Norm*, $\phi$, measures whether voters punish parties that do not follow the norm *in cases in which the two most voted councilors are of the same gender*, and the coefficient on *Norm∗GenderDiff*, $\xi$, captures whether punishment is different when the top two councilors are of different gender. In particular, if voters drive the gender differential effects, then $\xi$ should be negative.

The results are shown in table 6. In column (1), which does not include party fixed effects, we see that parties that follow the norm gain 7.5 p.p. more votes in the next election compared with those that break the norm. In column (2), where we add party fixed effects, the effect remains similar (6.5 p.p.) and significant at the 5% level. Column (3) studies gender differences in the enforcement of the norm. The coefficient on $Norm$ is 7.3 p.p., which indicates that, when a party follows the norm and the two most voted councilors were of the same gender, in the next election the party gains 7.3 p.p. in vote share. The coefficient on *Norm∗GenderDiff* is -7.8 p.p., indicating that the gain from following the norm disappears completely when the most voted was female and the second was male. Column (4) shows that these results are robust to including party fixed effects. Finally, column (5) performs a placebo test, in which we show that there is no effect of the norm on the lagged outcome variable—that is, parties that follow the norm have obtained a similar vote share in the time $t$ election than those who break the norm. Hence, these results support that voters enforce the norm and its gender differences.

## 5.3 Effect of the Electoral System by Previous Exposure to a Female Mayor

The model predicts that, in the absence of voter and party biases, the two electoral systems should result in the same share of female candidates, councilors, and mayors. Here we test this prediction. As shown in section 5.1, women attract, on average, fewer votes than men, but this is not the case in municipalities that have been exposed to a female mayor, suggesting that there is no (voter or party) bias against women in these municipalities. Hence, we split the sample by exposure to a female mayor and run equation (10) separately in the two subsamples. Under the model's assumptions, we should see that the positive effect of CL on the share of female candidates, councilors, and mayors is not present in the subsample of exposed municipalities.

The results are in table 7 (also see figure 6). For candidates (panel A), CL increases the share of female candidates by between 3.5 p.p. and 4.5 p.p. in non-exposed municipalities. In exposed municipalities, by contrast, the point estimates of CL are smaller and non-significant in most

[55] In OL, voters do not vote directly for a party, so it is not obvious how to aggregate votes by party. We follow the formula established in the Spanish Electoral Law to measure the strength of parties in municipal elections. According to this formula, which was created to allocate seats to parties in province governments, votes for a party are obtained as the average of the votes obtained by the individual candidates of that party, up to a maximum of four.

specifications. We can reject that the coefficients in exposed and non-exposed municipalities are equal at the 1% level. For councilors (panel B), CL increases women's representation by between 3.7 p.p. and 4.3 p.p. in non-exposed municipalities. In exposed municipalities, there is no significant effect of the electoral system. We can reject the null that the coefficients in exposed and non-exposed coefficients are equal, at the 5% level. Finally, the point estimates for mayors (panel C) are considerably larger in exposed municipalities, contrary to what we expected. There are two reasons that can explain this apparently surprising result. First, these coefficients are very noisily estimated, so it is hard to draw strong conclusions. In fact, the coefficients in the subsample of exposed municipalities, although quantitatively large, are non-significant in most specifications, and we cannot reject the null that the estimates are equal in the exposed and non-exposed subsamples. Second, the percent (as opposed to percentage points) effect of CL is substantially lower in exposed municipalities compared with non-exposed municipalities.[56]

# 6 Other Possible Mechanisms

In this section, we discuss three alternative explanations to our findings. First, the CL system could increase the share of women because of its larger district magnitude. Our model relies on the different ballot structure of the two electoral systems. However, there is also a difference in the district magnitude: seven councilors are elected in CL, and five in OL. It has been argued that larger district magnitudes lead to more women's representation (Matland and Brown 1992). A possible reason is that more seats make "ticket balancing" easier. Political parties want to put some men at the top of the list, and then balance the list with females at the bottom (if there is room for them). Therefore, the longer the list is (the larger the district magnitude) the easier it is for parties to balance their tickets. As the CL system has the larger district magnitude, it should make it easier for parties to include more women. However, note that, while this mechanism might explain the effect on candidates or councilors, it cannot account for the effect on mayors. Furthermore, this mechanism cannot explain why the effect of CL disappears in municipalities with low or no bias against women. Finally, and more importantly, we exploit the fact that, at 1,000 inhabitants, there is another threshold in which there is a change in district magnitude but not in the ballot structure. Municipalities with fewer than 1,000 inhabitants (and more than 250) elect seven councilors, while those with 1,000 or more inhabitants elect nine councilors—hence, they both use the CL system but with different district magnitudes. We perform an analysis in this threshold using the same empirical strategy as for the 250-inhabitant threshold. The results from these tests, displayed in table 8 and figure 7, reveal no effect of the district magnitude on the share of candidates, councilors, or mayors.[57]

Second, the effects might be driven by a *supply* component. The literature has discussed three main reasons behind women's underrepresentation in politics: voter bias, party bias, and supply reasons, i.e. women's (un)willingness to participate in politics. While the results in our paper suggest that voter bias is more relevant than party bias, we cannot say anything about the strength of voter bias relative to the supply component. In fact, it is very likely that there is a supply component at play. In particular, note that the share of women is substantially below 50%, and it is hard for voter (or party) bias to explain such low levels. A different question is whether a supply mechanism can drive *the difference* in women's representation between the two electoral systems. For this to happen, it is not sufficient that there is a supply component at play, but this supply component must also differ by electoral system. While we cannot rule out this possibility, we believe that it cannot drive the entirety of the effect. First of all, there is no obvious reason why women's

[56] For example, based on the IK estimation of column (1), the percent effect is 21% (=11.3/54.4) in exposed municipalities, and 56% (=3.5/6.2) in non-exposed municipalities.

[57] Figure A4 shows that there is no manipulation of the running variable around this threshold (p-value=0.63), and table A1 shows that covariates are balanced.

willingness to participate should be different in OL and CL.[58] Furthermore, a supply component cannot directly explain why the effect disappears in the subsample of exposed municipalities.

Third, the two electoral systems might lead to different party structures, and this in turn have an impact on women's presence. For example, if one electoral system favors left-wing parties, and left-wing parties include more women in their lists, this system could end up with a higher share of female candidates. To shed light on this mechanism, we conduct two tests. First, we study whether the electoral system affects which party the mayor belongs to. We estimate equation (10) with the party of the mayor as the dependent variable. Table A2 and figure A5 show the results of these tests. There is no evidence that the electoral system affects which party is in office: the point estimates of CL are very close to zero and are not statistically significant. Second, we test whether the share of female candidates, councilors, and mayors is different in the two systems, *for a given party*. The results are shown in table A3 and figure A6. There is no heterogeneity in the effect for candidates and councilors: the two main parties, the PP and the PSOE, both include a higher share of women candidates, and have a higher share of women among their councilors in CL. For mayors, the point estimate is much larger for the PP, but the results are imprecise and not statistically significant for any party.

# 7 Conclusion

We have developed a flexible model to estimate the effects of voter and party gender biases on candidate choice and mayor selection, and thus on observed female representation across electoral systems. The model assumes that candidates are in limited supply, and that voters care about the gender and competence of candidates and about party ideology. Preferences are aggregated through probabilistic voting. Parties choose candidates to maximize their probability of winning an election aware of voter bias, and possibly subject to party bias. The model has stark predictions on how voter and party biases affect female representation across two electoral systems, an OL system in which voters choose candidates from all lists, and a CL system in which voters choose an entire list of candidates.

We have used the context of Spanish municipal elections to test the model predictions. Given that the electoral system is determined by population size as mandated by national law, RD provides credible estimates of causal effects. Our regressions show that the CL system increases the share of female candidates and councilors by between 2 and 3 p.p., and the share of female mayors by between 3.5 and 6 p.p. Our model explains these results as mostly driven by average *voter* bias against women. We have provided additional evidence that supports the mechanism of the model: women attract fewer votes than men; there is a norm that the most voted should be appointed mayor in the OL system, and there are gender differences in this norm, which are enforced by voters; and the differences in women's representation between CL and OL are not present in a subsample of municipalities in which there is arguably less bias against women.

Our paper studies the effect of the electoral system on both directly- and indirectly-elected offices, i.e., councilors and mayors, respectively. Much of the work on women's representation is about elected offices, but the evidence on selected executive posts is scarce and mixed. For example, regarding the effects of quotas, O'Brien and Rickne (2016) find that quotas can have an acceleration effect on women's representation in leadership positions, but Bagues and Campa (2017a) find that quotas are successful in increasing women's representation in directly elected office, but fail to lift the barriers to indirectly-elected, leadership positions. This is important because these indirectly-elected offices hold, in many cases, most of the power. Our theoretical and empirical results show that the electoral system may have a substantial impact in increasing women's representation in both directly- and indirectly-elected offices.

[58] One such possible reason is that the OL system leads to a more aggressive or competitive campaign. If women shy away from competition, then women might be less interested in running in OL. But, if this were the case, we should see no difference in the share of female mayors between the two electoral systems: given that, in CL, the mayor candidate has more visibility and campaign responsibilities, women should shy away from being mayor candidates in this system.

Throughout the paper, we have shown that an important driver of the results is that voters care about the gender and quality of the top-listed candidate in CL, while they care about the gender and quality of all candidates in OL. In the context of our paper, one natural explanation is that mayors have most of the power and, in CL, only top-listed candidates can become mayor, while, in OL, any councilor can be appointed mayor. An interesting question is whether voters would behave in a similar way if the council itself had the executive power and there was no mayor selection or if, alternatively, the council had to select a mayor but any councilor could be appointed mayor in both CL and OL. Given that a growing amount of evidence suggests that voter fatigue and inattention are relevant aspects of voter behavior (Augenblick and Nicholson 2015), voters in CL might still pay substantially more attention to the traits of the top-listed candidate than to those of other candidates, i.e. voters, who need to pick one entire party-list, may use the characteristics of the top-listed candidate as an indicator of the quality of the whole list. In OL, by contrast, voters are forced to make a decision for any individual vote they cast, which may trigger them to pay attention to the characteristics of each individual candidate that they have to choose.

While RD designs are well known for providing a credible *internal* estimation of causal effects, they can be less conclusive with respect to the *external* validity of the findings, i.e. what the effects would be in other contexts. In this regard, note that, if the CL and OL systems were implemented in other settings, our theoretical model would yield predictions that would depend on voter and party biases in those settings. Furthermore, the model can be adapted to study the effect of other electoral systems on female representation. For example, consider the election of several, equally powerful, representatives. Under a multi-member, proportional-representation system, when deciding which party to vote for, voters would have to form expectations on the probability that each candidate of a given list is elected councilor. Under a first-past-the-post electoral system, there would be one election per district. Of particular interest in this case would be the effect of geographic restrictions on candidates: either candidates have to be from the district they represent, or from the country at large. We intend to explore this in future work.

# References


Anagol, S. and T. Fujiwara (2016). The runner-up effect. *Journal of Political Economy 124*(4), 927–991.

Atkinson, A. B., T. Piketty, and E. Saez (2011). Top incomes in the long run of history. *Journal of Economic Literature 49*(1), 3–71.

Augenblick, N. and S. Nicholson (2015). Ballot position, choice fatigue, and voter behaviour. *Review of Economic Studies 83*(2), 460–480.

Bagues, M. and P. Campa (2017a). Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design. *Unpublished Manuscript*.

Bagues, M. and P. Campa (2017b). Women and power: Unpopular, unwilling, or held back? Comment. *Unpublished Manuscript*.

Bartels, C. and M. Metzing (2017). An integrated approach for top-corrected Ginis. *Working Paper, IZA DP No. 10573*.

Beaman, L., R. Chattopadhyay, E. Duflo, R. Pande, and P. Topalova (2009). Powerful women: Does exposure reduce bias? *Quarterly Journal of Economics 124*(4), 1497–1540.

Beaman, L., E. Duflo, R. Pande, and P. Topalova (2012). Female leadership raises aspirations and educational attainment for girls: A policy experiment in India. *Science 335*(6068), 582–586.

Becker, G. S. (1957). *The economics of discrimination.* The University of Chicago Press.

Besley, T. and S. Coate (1997). An economic model of representative democracy. *Quarterly Journal of Economics 112*(1), 85–114.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Casas-Arce, P. and A. Saiz (2015). Women and power: Unwilling, ineffective, or held back? *Journal of Political Economy 123*(3), 641–669.

Cattaneo, M. D., B. R. Frandsen, and R. Titiunik (2015). Randomization inference in the regression discontinuity design: An application to party advantages in the us senate. *Journal of Causal Inference 3*(1), 1–24.

Cattaneo, M. D., M. Jansson, and X. Ma (2017). Simple local polynomial density estimators. *Unpublished Manuscript*.

Chattopadhyay, R. and E. Duflo (2004). Women as policy makers: Evidence from a randomized policy experiment in India. *Econometrica 72*(5), 1409–1443.

Dahl, G., A. Kotsadam, and D.-O. Rooth (2018). Does integration change gender attitudes? The effect of randomly assigning women to traditionally male teams. *NBER Working Paper 24351*.

De Paola, M., V. Scoppa, and R. Lombardo (2010). Can gender quotas break down negative stereotypes? Evidence from changes in electoral rules. *Journal of Public Economics 94*(5), 344–353.

Egan, M. L., G. Matvos, and A. Seru (2017). When Harry fired Sally: The double standard in punishing misconduct. *NBER Working Paper 23242*.

Esteve-Volart, B. and M. Bagues (2012). Are women pawns in the political game? Evidence from elections to the Spanish Senate. *Journal of Public Economics 96*(3), 387–399.

Folke, O., T. Persson, and J. Rickne (2016). The primary effect: Preference votes and political promotions. *American Political Science Review 110*(3), 559–578.

Foremny, D. and A. Solé-Ollé (2016). Who's coming to the rescue? Revenue-sharing slumps and implicit bailouts during the Great Recession. *Unpublished Manuscript*.

Fortin-Rittberger, J. and B. Rittberger (2014). Do electoral rules matter? Explaining national differences in women's representation in the European Parliament. *European Union Politics 15*(4), 496–520.

Fréchette, G. R., F. Maniquet, and M. Morelli (2008). Incumbents' interests and gender quotas. *American Journal of Political Science 52*(4), 891–909.

Fujiwara, T. and C. Sanz (2018). Norms in bargaining: Evidence from government formation. *Unpublished Manuscript*.

Galasso, V. and T. Nannicini (2017). Political selection under alternative electoral rules. *Public Choice 171*(3-4), 257–281.

Imbens, G. and K. Kalyanaraman (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies 79*(3), 933–959.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*(2), 281–355.

Lindbeck, A. and J. Weibull (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice 52*(3), 273–297.

Magre-Ferran, J. and X. Bertrana-Horta (2005). Municipal presidentialism and democratic consolidation in Spain. *R. Berg, N. Rao (red.), Transforming Local Political Leadership, New York: Palgrave Macmillan*, 73–84.

Matland, R. E. and D. D. Brown (1992). District magnitude's effect on female representation in US state legislatures. *Legislative Studies Quarterly*, 469–492.

O'Brien, D. Z. and J. Rickne (2016). Gender quotas and women's political leadership. *American Political Science Review 110*(1), 112–126.

Osborne, M. J. and A. Slivinski (1996). A model of political competition with citizen-candidates. *Quarterly Journal of Economics 111*(1), 65–96.

Pande, R. and D. Ford (2011). Gender quotas and female leadership: A review. background paper for the World Development Report on gender.

Persson, T. and G. E. Tabellini (2002). *Political economics: Explaining economic policy*. MIT press.

Roberts, A., J. Seawright, and J. Cyr (2013). Do electoral laws affect womens representation? *Comparative Political Studies 46*(12), 1555–1581.

Sanz, C. (2017a). Direct democracy and government size: Evidence from Spain. *Working Paper 1709, Bank of Spain*.

Sanz, C. (2017b). The effect of electoral systems on voter turnout: Evidence from a natural experiment. *Political Science Research and Methods 5*(4), 689–710.

Sarsons, H. (2017). Interpreting signals in the labor market: Evidence from medical referrals. *Unpublished Manuscript*.

Schmidt, G. D. (2009). The election of women in list PR systems: Testing the conventional wisdom. *Electoral Studies 28*(2), 190–203.

Thames, F. C. and M. S. Williams (2010). Incentives for personal votes and womens representation in legislatures. *Comparative Political Studies 43*(12), 1575–1600.

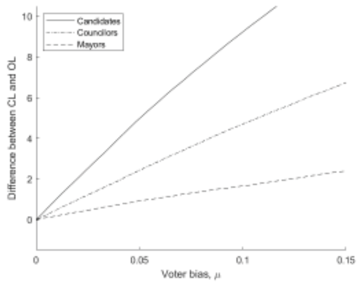
Valdini, M. E. (2012). A deterrent to diversity: The conditional effect of electoral rules on the nomination of women candidates. *Electoral Studies 31*(4), 740–749.

Vallés, J. M. and Q. Brugué (2001). El gobierno local. *Política y Gobierno en España*, 103–121.

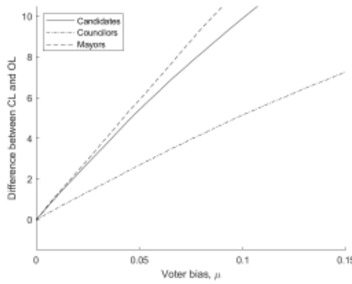
# Figures

Figure 1: CL vs. OL: Model Simulations

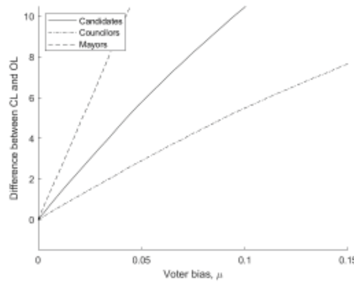
A. $\mu > 0, \nu = 0$




(a) Strong Norm

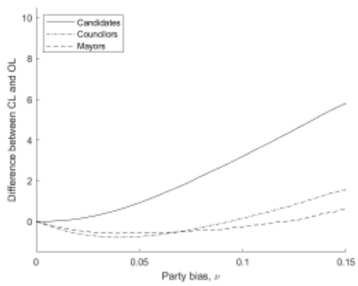



(b) Moderate Norm

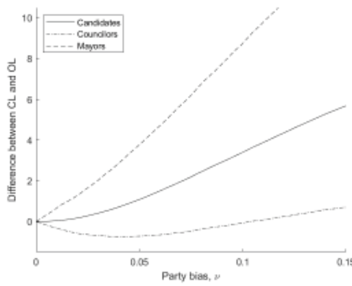



(c) Weak Norm

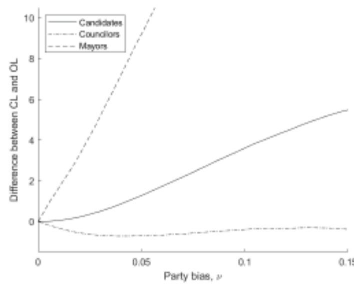
B. $\mu = 0, \nu > 0$




(d) Strong Norm



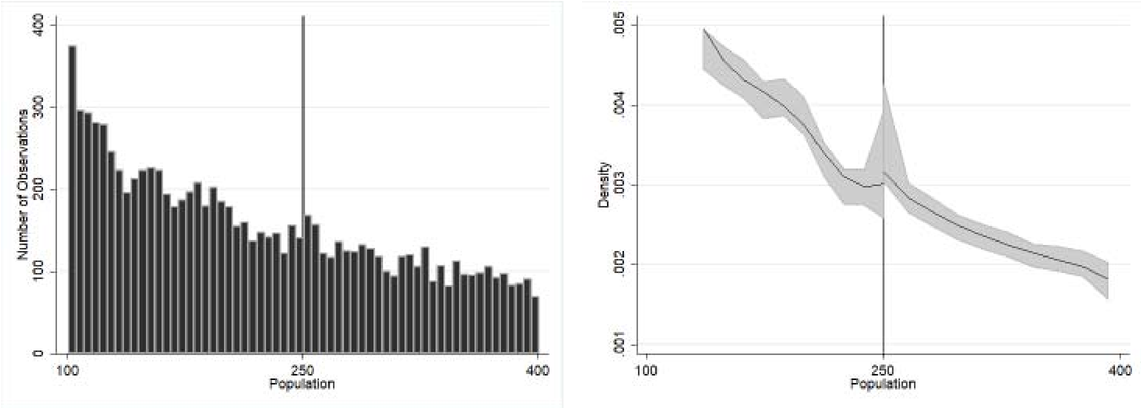

(e) Moderate Norm




(f) Weak Norm

The lines show the expected difference in the share of female candidates, councilors, or mayors, between the CL and the OL systems. Results are based on 1 million simulations of the theoretical model using a Pareto distribution for the competence of candidates. Panel A shows the case of voter bias and no party bias. Panel B shows the case of party bias and no voter bias. For the strong-norm cases, we choose the parameters $\hat{\beta}$ and $\hat{\alpha}$ such that, when evaluated at a level of voter bias of 0.05, the probability that the norm is not followed when the most voted candidate is female is around 3%. For the moderate- and weak-norm cases, we target probabilities that the norm is not followed when the most voted candidate is female of 15% and 30%, respectively.

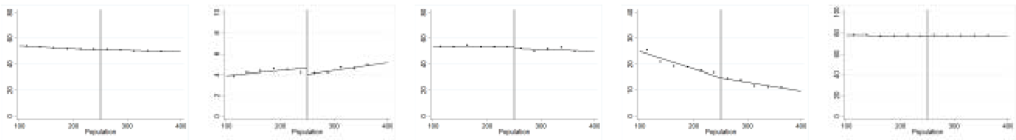
Figure 2: CL vs. OL: Test of Manipulation of the Running Variable



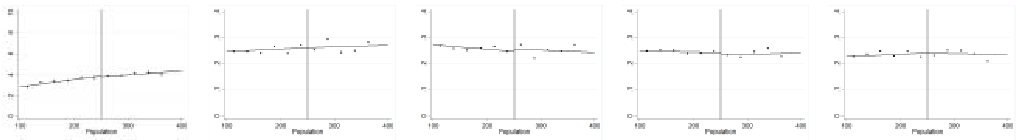

(a) Histogram (b) Test of manipulation

Figure (a) shows the histogram of population sizes. The unit of observation is a municipality-election. Bins are 25-inhabitant wide. Figure (b) shows the manipulation test by Cattaneo, Jansson, and Ma (2017): p-value = 0.39.

Figure 3: CL vs. OL: Covariate Smoothness around the Threshold



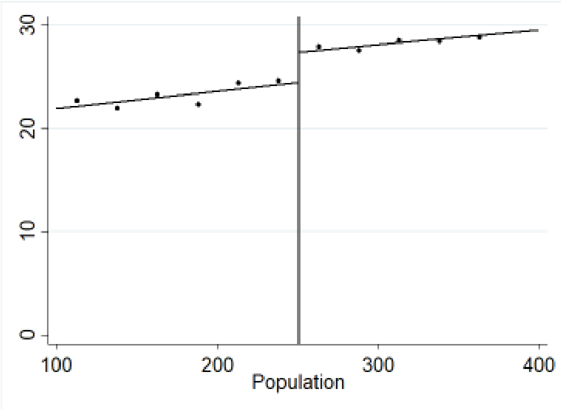
(a) Average age (years) (b) Foreigners (%) (c) EU foreigners (%) (d) Ideology (%) (e) Turnout (%)



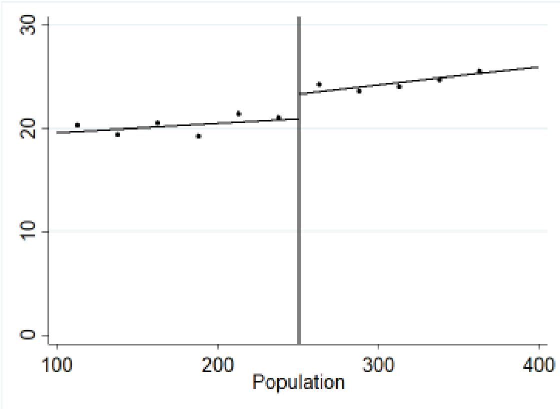
(f) Unemployment (%) (g) $\mathbb{1}_{year=2003}$ (h) $\mathbb{1}_{year=2007}$ (i) $\mathbb{1}_{year=2011}$ (j) $\mathbb{1}_{year=2015}$

The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 25 p.p.-wide bins of population (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

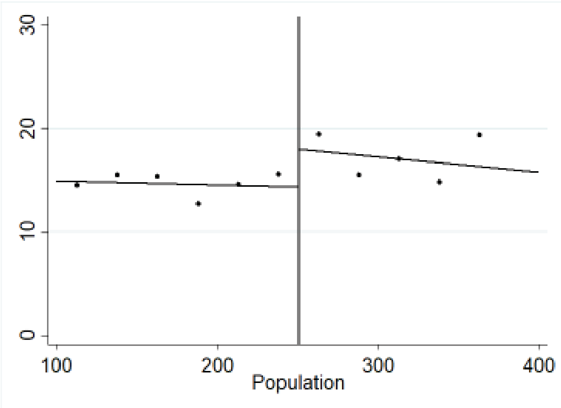
Figure 4: CL vs. OL: Effect on the Share of Female Candidates, Councilors, and Mayors




(a) Share of female candidates (%)



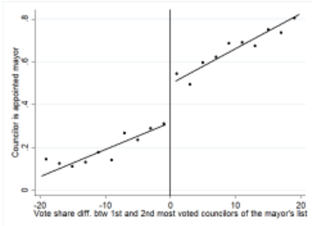

(b) Share of female councilors (%)



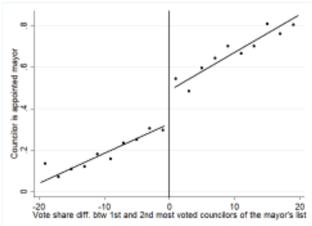

(c) Share of female mayors (%)

The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 25 p.p.-wide bins of population (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

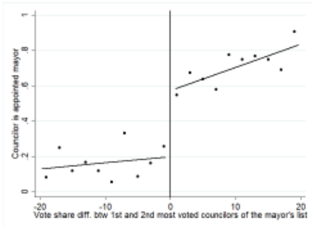
Figure 5: Mayor Selection in OL: Norm of the Most Voted, and Gender Differences in the Norm



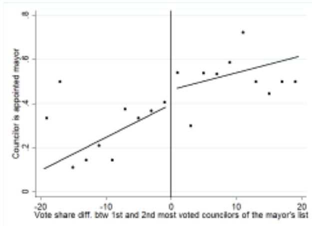

(a) All sample




(b) Same gender




(c) First male, second female

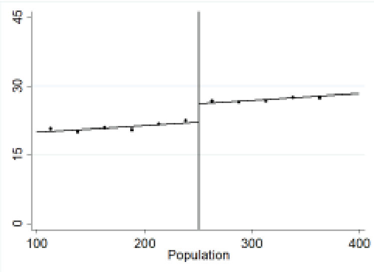



(d) First female, second male

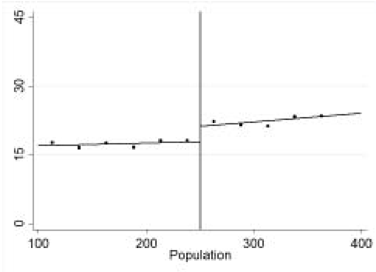
The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 2 p.p.-wide bins of the vote share difference of the top two councilors of the mayor's party (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

Figure 6: CL vs OL: Effects by Previous Exposure to a Female Mayor

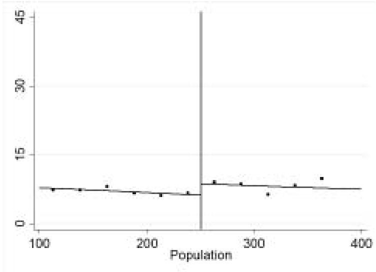
A. Non-Exposed Municipalities



(a) Share of female candidates (%)

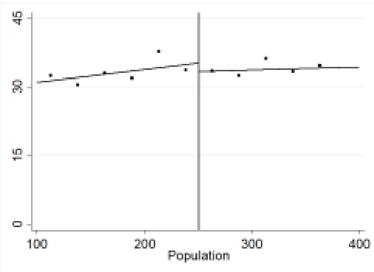


(b) Share of female councilors (%)

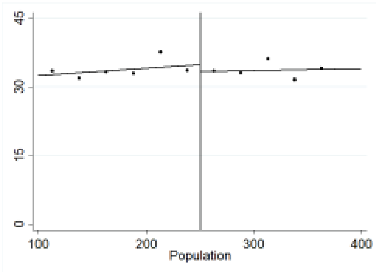


(c) Share of female mayors (%)

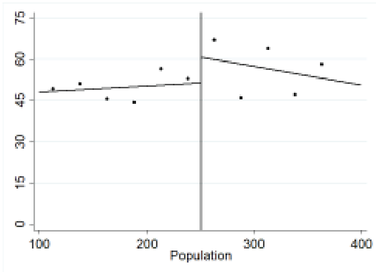
B. Exposed Municipalities



(d) Share of female candidates (%)



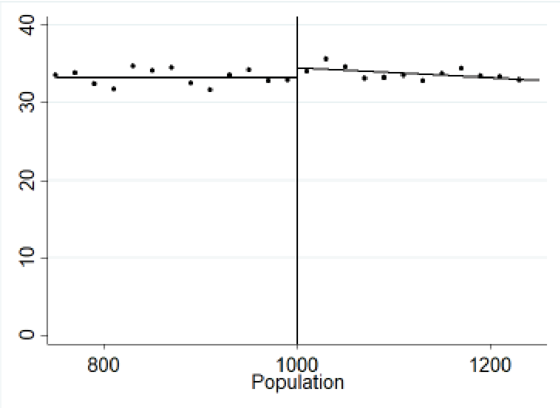
(e) Share of female councilors (%)



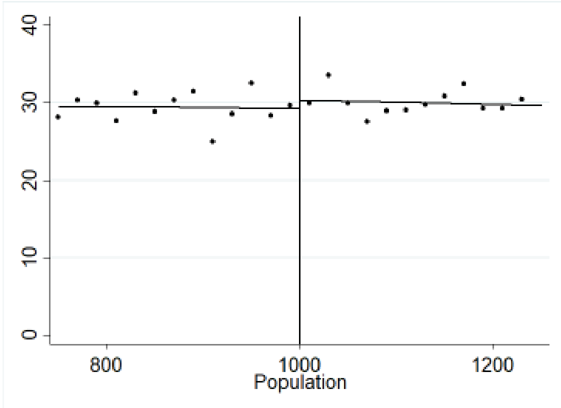
(f) Share of female mayors (%)

The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 25 p.p.-wide bins of population (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

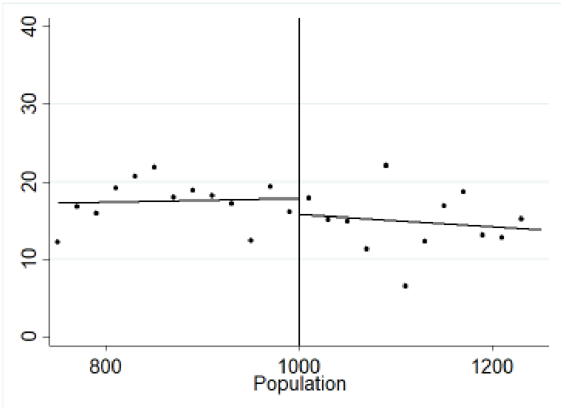
Figure 7: The Effect of District Magnitude: 1,000-Inhabitant Threshold




(a) Share of female candidates (%)




(b) Share of female councilors (%)




(c) Share of female mayors (%)

The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 25 p.p.-wide bins of population (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

## Tables

Table 1: Average Shares of Female Candidates, Councilors, and Mayors

Panel A: By Year

| Year | Sh Fem Candidates | Sh Fem Councilors | Sh Fem Mayors |
|---|---|---|---|
| 2003 | 20.8 | 17.3 | 12.8 |
| | (2,443) | (2,443) | (2,440) |
| 2007 | 24.9 | 21.4 | 15.1 |
| | (2,446) | (2,446) | (2,445) |
| 2011 | 26.5 | 23.2 | 16.3 |
| | (2,307) | (2,307) | (2,302) |
| 2015 | 27.7 | 25.5 | 18.1 |
| | (2,236) | (2,236) | (2,232) |
| Average | 24.9 | 21.7 | 15.3 |
| | (N=9,432) | (N=9,432) | (N=9,419) |

Panel A: By Electoral System

| Year | Sh Fem Candidates | Sh Fem Councilors | Sh Fem Mayors |
|---|---|---|---|
| CL | 28.4 | 24.6 | 17.0 |
| | (3,296) | (3,296) | (3,291) |
| OL | 23.0 | 20.2 | 14.7 |
| | (6,136) | (6,136) | (6,128) |
| Average | 24.9 | 21.7 | 15.3 |
| | (N=9,432) | (N=9,432) | (N=9,419) |

The unit of observation is a municipality-election. Number of observations in parentheses.

Table 2: CL vs. OL: Covariate Smoothness around the Threshold

| Variable | (1)<br>Mean | (2)<br>IK Bandwidth | (3)<br>CCT Bandwidth | (4)<br>IK Results | (5)<br>CCT Results |
|---|---|---|---|---|---|
| Average Age (years) | 51.33 | 59 | 43 | 0.371 | 0.133 |
| | | [N = 3350] | [N = 2458] | (0.343) | (0.396) |
| Foreigners (%) | 4.097 | 60 | 52 | -0.0494 | 0.0437 |
| | | [N = 3466] | [N = 2920] | (0.343) | (0.353) |
| EU Foreigners (%) | 53.14 | 93 | 51 | -2.509 | -1.535 |
| | | [N = 4641] | [N = 2496] | (2.389) | (2.949) |
| Ideology (%) | 16.02 | 75 | 39 | -0.363 | -1.483 |
| | | [N = 4322] | [N = 2173] | (1.782) | (2.232) |
| Turnout (%) | 77.20 | 139 | 33 | 0.223 | 0.245 |
| | | [N = 8556] | [N = 1846] | (0.427) | (0.630) |
| Unemployment (%) | 3.751 | 65 | 29 | -0.0898 | -0.0419 |
| | | [N = 3691] | [N = 1679] | (0.205) | (0.270) |
| $\mathbb{1}_{year=2003}$ | 0.272 | 88 | 33 | -0.00231 | -0.0150 |
| | | [N = 5150] | [N = 1846] | (0.0233) | (0.0377) |
| $\mathbb{1}_{year=2007}$ | 0.258 | 98 | 37 | -0.00151 | 0.0458 |
| | | [N = 5716] | [N = 2057] | (0.0219) | (0.0420) |
| $\mathbb{1}_{year=2011}$ | 0.248 | 93 | 34 | -0.0188 | -0.00618 |
| | | [N = 5386] | [N = 1940] | (0.0250) | (0.0449) |
| $\mathbb{1}_{year=2015}$ | 0.229 | 88 | 39 | 0.00456 | 0.00135 |
| | | [N = 5078] | [N = 2232] | (0.0219) | (0.0331) |

Column (1) shows the mean of the variables. Columns (2) and (3) show the Imbens and Kalyanaraman (2012) and Calonico, Cattaneo, and Titiunik (2014) bandwidth values and number of observations for a placebo test that estimates the effect of CL on the corresponding variable. Columns (4) and (5) show the results for the placebo tests: each column is a separate local linear regression with a uniform kernel. Standard errors, clustered at both municipality and the running variable, are in parentheses.

Table 3: CL vs. OL: Effect on the Share of Female Candidates, Councilors, and Mayors

Panel A: Share of Female Candidates

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *CL* | 1.955** | 2.763** | 2.911*** | 2.510** | 3.154*** |
| | (0.987) | (1.092) | (0.765) | (1.027) | p-v = 0.000 |
| Constant | 25.67 | 24.49 | 24.48 | 25.39 | |
| | (0.860) | (1.032) | (0.611) | (0.889) | |
| Observations | 4256 | 2920 | 9432 | 9432 | 1746 |
| Bw Size | 73 | 51 | 150 | 150 | 31 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Panel B: Share of Female Councilors

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *CL* | 2.664** | 3.185** | 2.346** | 2.541* | 2.677*** |
| | (1.279) | (1.500) | (1.009) | (1.406) | p-v = 0.006 |
| Constant | 21.63 | 21.29 | 21.00 | 21.77 | |
| | (1.057) | (1.238) | (0.775) | (1.145) | |
| Observations | 4905 | 2870 | 9432 | 9432 | 1746 |
| Bw Size | 85 | 50 | 150 | 150 | 31 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Panel C: Share of Female Mayors

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *CL* | 3.718 | 6.324* | 3.652** | 4.581* | 3.479** |
| | (2.630) | (3.514) | (1.838) | (2.633) | p-v = 0.045 |
| Constant | 15.90 | 14.65 | 14.38 | 15.15 | |
| | (2.191) | (3.029) | (1.399) | (2.192) | |
| Observations | 3756 | 1844 | 9419 | 9419 | 1744 |
| Bw Size | 65 | 49 | 150 | 150 | 31 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Results from estimating $Outcome_{mt} = \alpha + \beta CL_{mt} + f(Pop_{mt} - 250) + u_{mt}$. Each column reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. The last column shows the results from Cattaneo, Frandsen, and Titiunik (2015)'s procedure. Standard errors clustered by municipality and population in parentheses.

Table 4: The Gender of Candidates and Vote Shares

Panel A: CL

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $FemCand_1$ | -3.809*** | -3.315*** | -4.907*** | 0.330 | -3.098*** |
| | (0.794) | (0.949) | (1.104) | (2.208) | (0.884) |
| $ExpCand_1$ | | 8.900*** | 9.183*** | 7.571*** | 6.973*** |
| | | (0.777) | (0.889) | (1.773) | (0.694) |
| $FemCand_{2-7}$ | | | | | -0.841** |
| | | | | | (0.327) |
| $ExpCand_{2-7}$ | | | | | 5.752*** |
| | | | | | (0.475) |
| Observations | 3806 | 2797 | 2211 | 586 | 2797 |
| Sample | All | All | Non-exposed | Exposed | All |

Panel B: OL

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $FemCand$ | -1.985*** | -1.261*** | -1.742*** | -0.0736 |
| | (0.330) | (0.319) | (0.354) | (0.623) |
| $ExpCand$ | | 9.451*** | 9.736*** | 7.514*** |
| | | (0.468) | (0.527) | (0.839) |
| Observations | 15129 | 15117 | 12252 | 2865 |
| Sample | All | All | Non-exposed | Exposed |

Panel A shows the results from estimating $VSh_{pmt} = \gamma(pt) + \gamma(m) + \theta_{fem}FemCandidate_{1,pmt} + \theta_{exp}ExpCandidate_{1,pmt} + \epsilon_{pmt}$ in CL elections. Column (5) adds the gender and the experience of other candidates on the party-list (i.e. candidates ranked from 2 to 7 on the party-list). Panel B shows the analogous OL regressions, where the unit of observation is an individual candidate. The sample is restricted to non-incumbent parties. Standard errors clustered by municipality in parentheses.

Table 5: Mayor Selection in OL: Norm of the Most Voted, and Gender Differences in the Norm

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| $First$ | 0.186** | 0.250*** | 0.298* | -0.144 |
| | (0.0765) | (0.0941) | (0.174) | (0.202) |
| Constant | 0.321 | 0.287 | 0.253 | 0.510 |
| | (0.0399) | (0.0492) | (0.0842) | (0.104) |
| Observations | 1704 | 1164 | 270 | 270 |
| Sample | All | Same | Male-Female | Female-Male |

Results from estimating $y_{cmt} = \theta_0 + \theta_1 First + f(x_{cmt}) + \epsilon_{cmt}$. The sample is restricted to the two most voted councilors of the party of the mayor, and to non-incumbent parties. Column (2) restricts the sample to observations in which the two most voted councilors were of the same gender; column (3), to those in which the most voted was male and the second was female; and column (4), to those in which the most voted was female and the second was male. All the regressions use the IK bandwidth (4.6 p.p.). Standard errors clustered by municipality in parentheses.

Table 6: Mayor Selection in OL: Voters Punish Parties that Break the Norm, and Gender Differences in Punishment

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| *Norm* | 7.499** | 6.493** | 7.330** | 6.083* | 1.268 |
| | (3.003) | (3.130) | (2.997) | (3.128) | (2.768) |
| *Norm∗GenderDiff* | | | -7.831** | -7.562** | |
| | | | (3.736) | (3.579) | |
| Observations | 708 | 708 | 708 | 708 | 610 |
| Outcome | t+4 | t+4 | t+4 | t+4 | t (placebo) |
| Party fixed effects | No | Yes | No | Yes | Yes |

Column (1) shows the results from fitting $VSh_{pmt+4} - VSh_{pmt} = \gamma(m) + \gamma(t) + \beta Norm_{mt} + \epsilon_{pmt}$. Columns (2), (4), and (5) add party fixed effects. Columns (3) and (4) add $Norm * GenderDiff_{mt}$ as a control. The outcome is $VSh_{pmt} - VSh_{pmt-4}$ in column (5). The sample is restricted to non-incumbent parties. Standard errors clustered by municipality in parentheses.

Table 7: CL vs. OL: Effects by Previous Exposure to a Female Mayor

Panel A: Share of Female Candidates

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| EXPOSED | | | | | |
| $CL$ | -3.567* | 0.292 | -1.894 | -3.424 | -0.933 |
| | (2.058) | (2.169) | (1.649) | (2.149) | p-v = 0.587 |
| Constant | 25.67 | 24.49 | 24.48 | 25.39 | |
| Observations | 756 | 535 | 1703 | 1703 | 324 |
| NON-EXPOSED | | | | | |
| $CL$ | 3.443*** | 3.700*** | 4.029*** | 3.819*** | 4.253*** |
| | (0.957) | (1.022) | (0.723) | (0.960) | p-v = 0.000 |
| Constant | 23.12 | 23.02 | 22.06 | 22.99 | |
| Observations | 3500 | 2385 | 7728 | 7728 | 1422 |
| Bw Size | 73 | 51 | 150 | 150 | 31 |

Panel B: Share of Female Councilors

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| EXPOSED | | | | | |
| $CL$ | -2.081 | 0.796 | -1.529 | -2.932 | -0.656 |
| | (2.910) | (3.275) | (2.236) | (3.144) | p-v = 0.757 |
| Constant | 35.26 | 32.20 | 34.91 | 35.73 | |
| (2.350) | | | | | |
| Observations | 881 | 524 | 1703 | 1703 | 324 |
| NON-EXPOSED | | | | | |
| $CL$ | 3.708*** | 4.249*** | 3.282*** | 3.732*** | 3.678*** |
| | (1.159) | (1.404) | (0.933) | (1.296) | p-v = 0.000 |
| Constant | 18.57 | 18.44 | 17.88 | 18.67 | |
| Observations | 4024 | 2346 | 7728 | 7728 | 1422 |
| Bw Size | 85 | 50 | 150 | 150 | 31 |

Panel C: Share of Female Mayors

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| EXPOSED | | | | | |
| $CL$ | 11.31 | 22.84** | 9.486 | 9.795 | 9.393** |
| | (8.162) | (9.020) | (5.912) | (8.171) | p-v = 0.045 |
| Constant | 54.37 | 50.18 | 51.38 | 56.28 | |
| Observations | 669 | 508 | 1699 | 1699 | 323 |
| NON-EXPOSED | | | | | |
| $CL$ | 3.533* | 3.072 | 2.561** | 3.206* | 2.971** |
| | (1.887) | (2.141) | (1.232) | (1.891) | p-v = 0.034 |
| Constant | 6.218 | 6.335 | 6.164 | 6.074 | |
| Observations | 3087 | 2299 | 7719 | 7719 | 1421 |
| Bw Size | 65 | 49 | 150 | 150 | 31 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Results from estimating $Outcome_{mt} = \alpha + \beta CL_{mt} + f(Pop_{mt} - 250) + u_{mt}$, separately for the subsamples of exposed and non-exposed municipalities. Each column reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. The last column shows the results from Cattaneo, Frandsen, and Titiunik (2015)'s procedure. Standard errors clustered by municipality and population in parentheses.

Table 8: The Effect of District Magnitude: 1,000-Inhabitant Threshold

Panel A: Share of Female Candidates

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $\mathbb{1}_{Pop_{mt}>1,000}$ | 1.162 | 1.764* | -0.541 | 0.722 | -1.044 |
| | (0.869) | (1.001) | (0.677) | (0.910) | p-v = 0.586 |
| Constant | 33.29 | 32.96 | 34.21 | 42.22 | |
| | (0.758) | (0.924) | (0.538) | (6.241) | |
| Observations | 2795 | 1683 | 6570 | 6570 | 148 |
| Bw Size | 201 | 51 | 150 | 150 | 10 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Panel B: Share of Female Councilors

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $\mathbb{1}_{Pop_{mt}>1,000}$ | 0.648 | 1.428 | -0.750 | 0.174 | -0.798 |
| | (1.235) | (1.317) | (0.920) | (1.215) | p-v = 0.772 |
| Constant | 29.50 | 29.20 | 30.52 | 39.74 | |
| | (1.039) | (1.104) | (0.736) | (8.428) | |
| Observations | 2341 | 1707 | 6570 | 6570 | 148 |
| Bw Size | 239 | 51 | 150 | 150 | 10 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Panel C: Share of Female Mayors

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $\mathbb{1}_{Pop_{mt}>1,000}$ | -0.911 | 0.755 | -4.120* | -2.069 | 2.582 |
| | (2.618) | (3.312) | (2.150) | (2.756) | p-v = 0.740 |
| Constant | 17.03 | 16.65 | 17.94 | 15.48 | |
| | (2.262) | (2.840) | (1.680) | (17.86) | |
| Observations | 2560 | 1329 | 6554 | 6554 | 148 |
| Bw Size | 219 | 51 | 150 | 150 | 10 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Results from estimating $Outcome_{mt} = \alpha + \beta \mathbb{1}_{Pop_{mt}>1,000} + f(Pop_{mt} - 1,000) + u_{mt}$. Each column reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. The last column shows the results from Cattaneo, Frandsen, and Titiunik (2015)'s procedure. Standard errors clustered by municipality and population in parentheses.

# Online Appendices

# A Probabilistic Voting

The microfoundations for the political part of the model laid out in section 3 are due to Lindbeck and Weibull (1987) and Persson and Tabellini (2002). For simplicity, there is only one group of voters in each municipality.[59] Since the candidates (and the voters) lack commitment, the election is driven by candidates' competence, voters' gender bias, and ideology: Voters prefer more competent candidates, as these would be more likely to implement good policies. At the same time, voters might have a bias against female candidates, thus preferring a less competent male candidate. Finally, which candidates a voter support depends not only on the candidates' competence (which is imperfectly observed by voters) and gender, but also on the relative "ideological" attachment of the voter to the candidate.

Consider the choice of two councilors among four competing candidates in a given municipality. Denote the candidates by $A1$, $A2$, $B1$, and $B2$, where it is understood that candidate $X1$ has higher competence net of voter bias than candidate $X2$. Voter $i$ supports candidates $A1$ and $A2$ if the voter's indirect utility from this pair of candidates exceeds the indirect utility from the other three possible combinations, by more than some threshold values. These threshold values have a voter-specific component, reflecting both the ideological attachment and imperfect observation of competence mentioned before, and an aggregate component.

The voter-specific component, denoted by $\sigma^i$, is drawn from a symmetric uniform distribution.[60] For simplicity, we assume that voter $i$ observes competence for both candidates of the same party with the same observational error.[61] The aggregate component, denoted by $\delta$, is drawn from a symmetric, uniform distribution, $\delta \sim U[-1/(2\psi), 1/(2\psi)]$. This component represents an aggregate shock to ideological attachment which is realized after the parties have chosen candidates. The sum of the two components represents the total bias of voter $i$ in favor of candidates from party $A$ in the current election.[62]

Let $v(\pi)$ denote the indirect utility function of any voter $i$ when pair of candidates $\pi \in \{A1A2, A1B1, B1A1, B1B2\}$ is in office.[63] We assume that

$$v^i(XY) = s^X - \mu \mathbb{1}_{g(X)=f} + \left(\sigma^i + \delta\right) \mathbb{1}_{X \in A} + \epsilon \left[s^Y - \mu \mathbb{1}_{g(Y)=f} + \left(\sigma^i + \delta\right) \mathbb{1}_{Y \in A}\right],$$

where $s^i$ is the competence of candidate $i$, $g(i)$ is its gender, $f$ for female, $m$ for male, $\mu$ measures voter gender bias, $\mathbb{1}_{X \in A}$ is an indicator for whether a candidate belongs to party $A$, and $\epsilon$ determines the relative importance of the second placed candidate in voters' preferences. The latter is used to distinguish between electoral systems. Given that the mayor is most influential, in CL systems, $\epsilon << 1$ and for simplicity we will derive results in the limit of $\epsilon = 0$. In contrast, in the OL system, $\epsilon = 1$ since a priori any elected councilor from the winning party can be selected as mayor.

For the cases of interest in the OL system, voters with strong attachment to party $B$ will vote $B1B2$. As ideology is dampened, voters vote $B1A1$, and for weak ideological attachment to party $A$ the vote switches to $A1B1$, with voters with strong ties to party $A$ voting $A1A2$. Since the party that gets two councilors wins the election, the election will be decided by the mass of votes $A1A2$ relative to $B1B2$. These are determined by the choice of voters that are indifferent on one hand between $A1A2$ and $A1B1$, and those that are indifferent between $B1B2$ and $B1A1$. These decisions

[59] The model can be extended to include two groups, say, male and female voters.

[60] With only one group of voters, there is no need to specify the distribution of voter specific ideology. With more than one group, the distributions must be specified, and the relative dispersion of ideologies determines the relative weights parties attach to groups when selecting candidates.

[61] Otherwise we would need to specify $\sigma^{i1}$ and $\sigma^{i2}$.

[62] Thus, a given value of $\sigma^i$ measures both a relative ideological attachment to party $A$, and the relative difference between the actual and observed competence of candidates from both parties. In traditional probabilistic voting models, the latter component is absent and $\sigma^i$ is a pure measure of ideology.

[63] Note that $A2B2$ and $B2A2$ are dominated, respectively, by $A1B1$ and $B1A1$ for all voters, and thus can be disregarded.

are determined by the comparisons of marginal voters identified by their specific components, $\sigma^{AB}$ and $\sigma^{BA}$ respectively,

$$\begin{aligned}
v^{AB}(A1A2) &= v^{AB}(A1B1) - \epsilon(\delta + \sigma^{AB}) \\
s^{A1} - \mu \mathbb{1}_{g(A1)=f} + \epsilon\left(s^{A2} - \mu \mathbb{1}_{g(A2)=f}\right) &= s^{A1} - \mu \mathbb{1}_{g(A1)=f} + \epsilon\left(s^{B1} - \mu \mathbb{1}_{g(B1)=f} - \delta - \sigma^{AB}\right) \\
v^{BA}(B1B2) &= v^{BA}(B1A1) + \epsilon(\delta + \sigma^{BA}) \\
s^{B1} - \mu \mathbb{1}_{g(B1)=f} + \epsilon\left(s^{B2} - \mu \mathbb{1}_{g(B2)=f}\right) &= s^{B1} - \mu \mathbb{1}_{g(B1)=f} + \epsilon\left(s^{A1} - \mu \mathbb{1}_{g(A1)=f} + \delta + \sigma^{BA}\right)
\end{aligned}$$

Note that, in the CL system, since $\epsilon = 0$, the comparison is just between $A1$ and $B1$, as in the main text.

In the OL system, the probability that party $A$ wins, $P^A$, is given by the probability that $\sigma^{AB} + \sigma^{BA} < 0$, i.e. the probability that there are more voters above $\sigma^{AB}$ than voters below $\sigma^{BA}$. From the expressions above, $P^A$ is given by the probability that $2\delta > s^{A1} + s^{A2} - s^{B1} - s^{B2} - \mu(\mathbb{1}_{g(A1)=f} + \mathbb{1}_{g(A2)=f} - \mathbb{1}_{g(B1)=f} - \mathbb{1}_{g(B2)=f})$. Given the assumption that $\delta$ is uniformly distributed, this implies

$$P^A = \frac{1}{2} + \psi \frac{1}{2}\left(s^{A1} + s^{A2} - s^{B1} - s^{B2} - \mu(\mathbb{1}_{g(A1)=f} + \mathbb{1}_{g(A2)=f} - \mathbb{1}_{g(B1)=f} - \mathbb{1}_{g(B2)=f})\right). \quad (15)$$

A similar derivation implies that, in the CL system, the probability that party $A$ wins the election is given by (3). In both cases, the probability that party $B$ wins is $1 - P^A$.

# B Proofs of Propositions 2 and 4

Proposition 2

From the comparison of (5) and (6) with (7) and (8), it follows that, if $\mu > 0$ and $\nu = 0$, then $P^{CL}(F = 2) > P^{OL}(F = 2)$ and $P^{CL}(F = 0) < P^{OL}(F = 0)$, and hence, $\mathrm{E}[F]^{CL} - E[F]^{OL} > 0$.[64] That is, in the presence of voter bias but no party bias, the CL system leads to more female candidates than the OL system.

By contrast, in the case of no voter gender bias ($\mu = 0$), if $\min[\nu, \beta] = \nu$, then $P^{CL}(F = 2) \leq P^{OL}(F = 2)$ and $P^{OL}(F = 0) \leq P^{CL}(F = 0)$, and thus a priori it is not possible to determine under which electoral system there will be more female candidates. But, if $\nu << 1$, then $P^{CL}(F = 2) \approx P^{OL}(F = 2)$ and $P^{OL}(F = 0) \approx P^{CL}(F = 0)$, and both system will lead to approximately the same share of female candidates.[65] When $\min[\nu, \beta] = \beta$, and $\nu >> E[f_1] - E[f_2]$ (such that we can approximate $\min[f_2, f_1 - \nu] = f_1 - \nu$), $P^{OL}(F = 2) = P(m_1 < f_2 - 2\nu + 2\beta)$ and $P^{OL}(F = 0) = P(f_1 < m_2) + P(m_1 < f_1 < m_2 + 2\beta)$, while $P^{CL}(F = 2) = P(m_1 < f_1 - \nu)$ and $P^{CL}(F = 0) = P(f_1 < m_2)$. Thus, $P^{OL}(F = 0) \approx P^{CL}(F = 0)$, while $P^{OL}(F = 2) < P^{CL}(F = 2)$, and there will be more female candidates in the CL system.

Proposition 4

If there is voter bias but no party bias, and $\beta(1 - \alpha\mu) > \nu = 0$, then the share of female mayors should be approximately the same across electoral systems. The reason for this is that if a party chooses to have both a male and a female candidate, the probability that the female candidate receives more votes than the male candidate in the OL system is also $P(f_1 > m_1 + \mu) = P^{CL}(g(ma(j)) = f)$.[66] If there is party bias but no voter bias, the analysis is simpler since the probability of winning the election will not be affected by candidates' gender. In this case, while $\nu < \beta$ we should see more female mayors under the OL system. The reason for this is that if a party chooses to have both male and female candidates, the probability that the female candidate

[64] This follows since in this case $P^{CL}(F = 0) = P(f_1 < m_2) < P(f_1 < m_2 + \mu) = P^{OL}(F = 0)$, and $P^{CL}(F = 2) = P(m_1 < \min[f_2, f_1 - \mu]) > P(m_1 < f_2 - \mu) = P^{OL}(F = 2)$.

[65] This follows since in this case $P^{CL}(F = 2) = P(m_1 < \min[f_2, f_1 - \nu]) \approx P(m_1 < f_2) = P^{OL}(F = 2)$ and $P^{OL}(F = 0) = P(f_1 < m_2) + P(m_1 < f_1 < m_2 + \nu) \approx P(f_1 < m_2) = P^{CL}(F = 0)$.

[66] If $\beta(1 - \alpha\mu) < 0$, the male candidate will always be selected as mayor in this case, and there would be less female mayors in the OL system.

receives more votes than the male candidate in the OL system is now $P(f_1 > m_1) > P(f_1 > m_1 + \nu) = P^{CL}(g(ma(j)) = f)$. But if $\nu > \beta$, then there will be less female mayors under the OL system, since now parties will always select the male runner-up over the female winner as mayor, and the only possibility for a female mayor is when both candidates are women. Thus, in this case

$$P^{OL}(g(ma(j)) = f) = P^{OL}(F = 2) < P^{CL}(F = 2) < P(f_1 > m_1 + \nu) = P^{CL}(g(ma(j)) = f).$$

# C Numerical Simulations

For the numerical simulations we assume that candidates' competence is drawn from a Pareto distribution.[67] Bartels and Metzing (2017) estimate the Pareto distribution parameter for Spain from 2001 and 2012. We take their average value of 1.8491, and we posit $\psi = 1$. We will consider biases between zero and 0.15. The model predicts that, in OL, if the most voted candidate is female and the second most voted candidate is male, then the party will select the male candidate as mayor as long as $\beta [1 - \mu\alpha] < \nu$, i.e. when the cost of deviating from the mayor selection norm is lower than the cost of choosing a female mayor. Whether this inequality is satisfied or not depends on the values that $\mu$, $\nu$, $\beta$, and $\alpha$ take for a given municipality in a given election. We make $\beta$ and $\alpha$ stochastic such that in equilibrium there is variation in the degree with which the norm is followed when the most voted candidate is female.[68]

Panel A of figure 1 considers the case of voter bias ($\mu$) when there is no party bias ($\nu = 0$).[69] In the first graph, we consider the case in which the norm is "strong", and hence, in the OL system, the probability that a woman is not appointed mayor after winning the popular vote is low.[70] In the second and third graphs, we consider the cases in which the norm is "intermediate" and "weak", respectively.[71]

We can see that the CL system leads to a higher share of female candidates than the OL system, and the gap is increasing in voter bias. This is consistent with proposition 2, which states that, if there is voter bias and no party bias, then the CL system increases the expected share of female candidates. Intuitively, in OL, parties avoid choosing female candidates *for the two spots* of the list, as they will attract fewer votes than male candidates. In CL, by contrast, voters only care about the gender of the top-listed candidate (as only he or she can become mayor), and hence parties will avoid choosing a female candidate *only for the first spot* of the list. Note that this effect is very similar in the three graphs of Panel A, indicating that the it does not depend on the strength of the norm.

The CL system also increases the share of female councilors, regardless of the strength of the norm. This effect, however, is smaller than the one for candidates. In OL, the share of female councilors is approximately the same as the share of female candidates, as parties internalized voter bias when choosing candidates. In CL, by contrast, the share of female councilors is lower than the share of female candidates (but still higher than the share of female councilors in OL). The reason is that the top-listed candidate is more likely to be male than female, while the converse is true for the second-listed candidate.[72] Given that only the top-listed candidate gets elected councilor of the party that loses the election, male candidates are, on average, more likely than female candidates to become councilors in CL.

[67]Empirical studies have documented that the upper tail distribution of income in most countries follows a Pareto distribution, see e.g. Atkinson, Piketty, and Saez (2011). Thus, we proceed under the assumption that income and competence are proportional.

[68]We assume uniform distributions for $\beta$ and $\frac{1}{\alpha}$: $\beta \sim U[0, \hat{\beta}]$ and $\frac{1}{\alpha} \sim U[0, \hat{\alpha}]$. Note that, if there is party bias and no voter bias, only $\beta$ matters, while, if there is voter bias and no party bias, only $\alpha$ matters.

[69]All the results are based on one million simulations.

[70]For this case, we choose the parameters $\hat{\beta}$ and $\hat{\alpha}$ such that, when evaluated at a level of voter bias of 0.05, the probability that the norm is not followed when the most voted candidate is female is around 3%.

[71]For these cases, we target probabilities that the norm is not followed when the most voted candidate is female of 15% and 30%, respectively.

[72]With some positive probability, parties choose a male candidate for the first spot that is not the most competent (due to voter bias), and, therefore, for the second spot, for which parties care only about competence, parties are left to choose from female potential candidates that are, on average, more competent than male ones.

Finally, the CL system also increases the share of female mayors. This effect is very sensitive to the strength of the norm. When the norm is strong, then parties follow the norm even when the first is female and the second is male. In this case, the difference between the two electoral system is very small, consistent with proposition 4. As the strength of the norm decreases, however, it becomes more likely that the OL norm is not followed when the most voted candidate is female and the second is male, reducing the share of female mayors in this system. Hence, the effect of CL is on the share of female mayors increases as the OL becomes weaker: when the norm is intermediate or weak, the CL increases the share of female mayors.

Panel B of figure 1 considers the case of party bias ($\nu$) when there is no voter bias ($\mu = 0$). As in Panel A, we consider three levels of strength in the norm. For low values of party bias, the two systems result in approximately the same share of female candidates, as stated in proposition 2. As party bias increases, CL results in relatively more female candidates. The intuition is as follows. In CL, parties know that, if they win the election, the top-listed candidate will become mayor. Given that parties are gender biased and therefore prefer to have a male than a female mayor, they are less likely to choose a woman as the top-listed candidate. However, they are more likely to choose a woman for the second spot of the list, as they only care about competence for that position, and women will be, on average, more competent among the remaining potential candidates (as explained in footnote 72). These two effects offset each other and the share of female candidates is always close to 50%. In OL, there are two forces at play. On the one hand, parties have a preference to choose male candidates, as that increases the probability of eventually having a male mayor. On the other hand, voters will vote for the most competent candidates, and will seek to enforce the norm that the most voted candidate becomes the mayor. Hence, choosing (less competent) male candidates instead of (more competent) female ones makes the party lose votes. As party bias increases, the former effect dominates and fewer female candidates are chosen in OL. As was the case with voter bias, note that changes in the strength of the norm have only minor effects on the share of female candidates (or councilors) in the case of party bias.

For councilors, and moderate values of party bias, the CL system leads to *fewer* female councilors than the OL system, consistent with proposition 3. The intuition behind this result is that, as explained, parties are less likely to put a woman at the top of the list in CL (given party bias, parties want to avoid that a woman becomes mayor if the party wins the election). For candidates, this effect is offset because the second spot of the list is more likely to be female. For councilors, however, this effect is not offset, as only the candidate at the top of the list is elected councilor from the party that loses the election, and that candidate is less likely to be female. Hence, in CL, the share of female councilors is below 50%. In OL, by contrast, voters elect the most competent candidates, and hence the share of female councilors is close to 50%.

Finally, the effect on the share of female mayors depends on the strength of the norm. If the norm is strong, then the OL system increases the share of female mayors. This finding, stated in proposition 4, mirrors that for councilors: in the OL system, the probability that a woman is the most voted candidate of the winning party is approximately 50% for low values of party bias. In contrast, in the CL system, a small party bias might affect the ordering of candidates within the lists, leading to a lower probability that a female candidate becomes mayor. If the norm is weak or intermediate, CL results in a higher share of female mayors than OL. As happened in the case of voter bias, as the strength of the norm decreases, it becomes more likely that the OL norm is not followed when the most voted candidate is female, reducing the share of female mayors in this system. Hence, the difference in the share of female mayors between CL and OL increases as the norm weakens.

## D Testing Gender Differences in the Norm

We formally test for gender differences in the norm using the following specification:

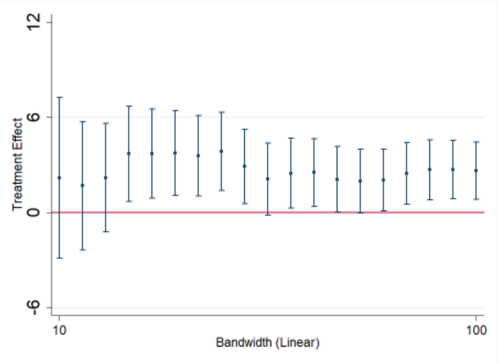
$$\begin{aligned} y_{cmt} = \theta_0 + \theta_1 \cdot First_{cmt} + g_0(x_{cmt}) + g_1(x_{cmt}) \cdot First_{cmt} + \theta_2 \cdot First_{cmt} \cdot Female1_{mt} + \theta_3 Female1_{mt} \\ +g_1(x_{cmt}) \cdot First_{cmt} \cdot Female1_{mt} + g_1(x_{cmt}) \cdot Female1_{mt} + \theta_4 \cdot First_{cmt} \cdot Female2_{mt} \\ +\theta_5 Female2_{mt} + g_1(x_{cmt}) \cdot First_{cmt} \cdot Female2_{mt} + g_1(x_{cmt}) \cdot Female2_{mt} + \epsilon_{cmt}, \end{aligned}$$

where $Female1_{mt}$ is a dummy that indicates whether the most voted councilor of the mayor's list was female, and $Female2_{mt}$ indicates whether the second most voted councilor of the mayor's list was female. Hence, these variables only vary by municipality-year, not by councilor. The interpretation of the coefficients is the following: $\theta_1$ is the probability that a male councilor who obtained the most votes becomes mayor, when the second is also male, relative to the probability that the second male councilor becomes mayor. That is, it measures the winner effect in a male-male environment. The rest of the coefficients are interpreted analogously. There is a bias against women if $\theta_2 < 0$, $\theta_3 > 0$, $\theta_4 > 0$, or $\theta_5 < 0$.

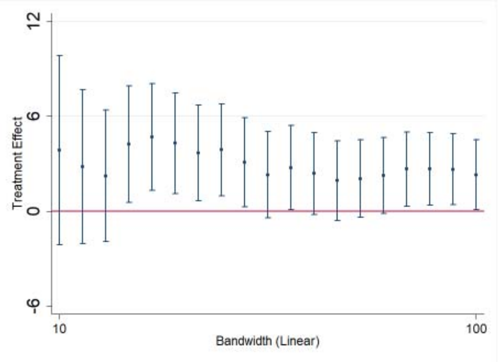
The results of this test are shown in column (1) of table A4. All of the coefficients are statistically significant and consistent with a bias against women in the norm. In the rest of the columns of table A4, we split the sample according to previous exposure to a female mayor. The results indicate that the bias is not present in exposed municipalities.

# Appendix Figures

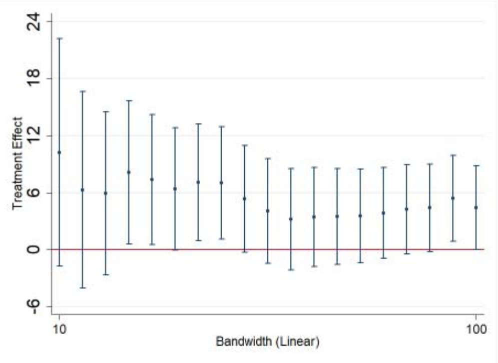
Figure A1: CL vs. OL: Effect on the Share of Female Candidates, Councilors, and Mayors—Robustness to Bandwidth Choice




(a) Share of female candidates (%)



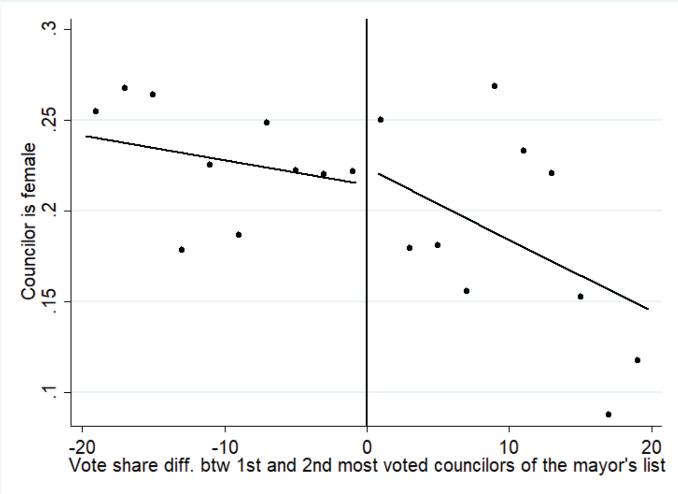

(b) Share of female councilors (%)




(c) Share of female mayors (%)

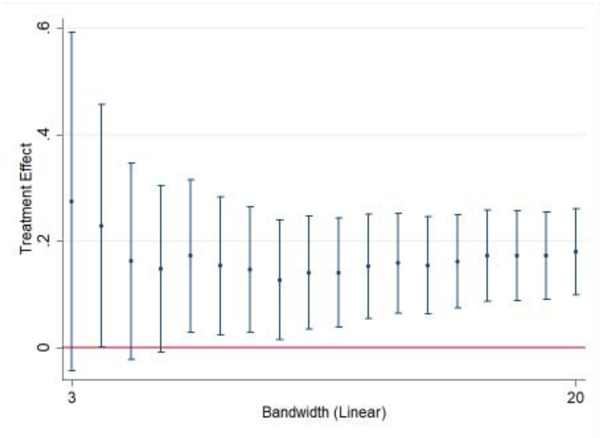
Circles represent the estimated treatment effect, using different bandwidth choices (x-axis), from $Outcome_{mt} = \alpha + \beta CL_{mt} + f(Pop_{mt} - 250) + u_{mt}$. Lines represent the 95% confidence interval (standard errors clustered by municipality and population).

Figure A2: Mayor Selection in OL: Placebo Test—"Effect" of the Most Voted on Being Female



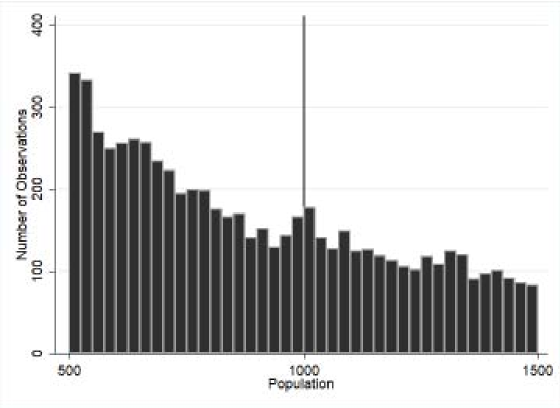

The unit of observation is a municipality-election. Circles represent the local averages of the variable (1 if female, 0 if male). Averages are calculated within 2 p.p.-wide bins of the vote share difference of the top two councilors of the mayor's party (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

Figure A3: Mayor Selection in OL: Norm of the Most Voted—Robustness to Bandwidth Choice

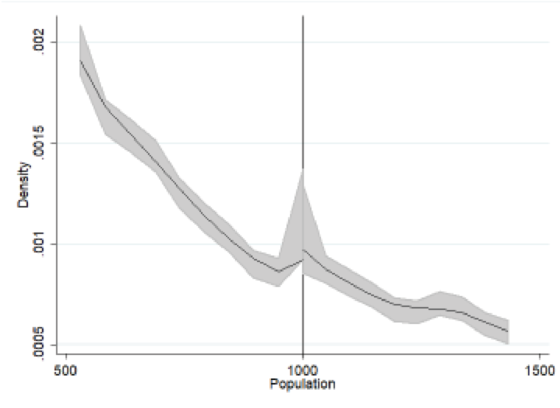



Circles represent the estimated treatment effect from fitting $y_{cmt} = \theta_0 + \theta_1 \cdot first + g_0(x_{cmt}) + g_1(x_{cmt}) \cdot first + \epsilon_{cmt}$, using different bandwidth choices (x-axis). Lines represent the 95% confidence interval (standard errors clustered by municipality).

Figure A4: 1,000-Inhabitant Threshold: Test of Manipulation of the Running Variable



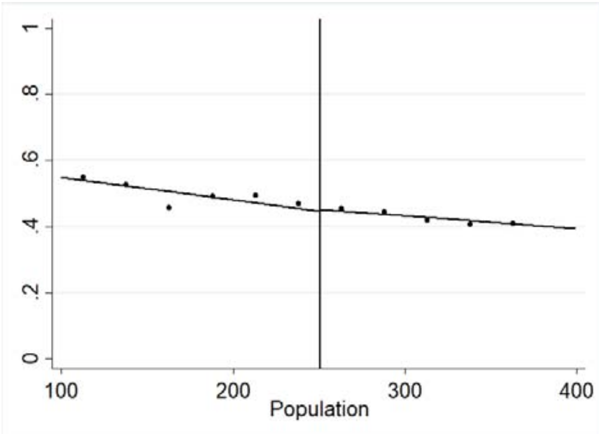

(a) Histogram



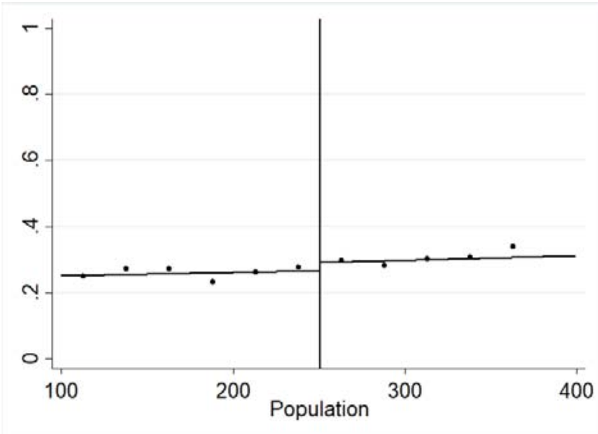

(b) Test of manipulation

Figure (a) shows the histogram of population sizes. The unit of observation is a municipality-election. Bins are 50-inhabitant wide. Figure (b) shows the manipulation test by Cattaneo, Jansson, and Ma (2017): p-value = 0.63.

Figure A5: CL vs OL: The Effect on the Party of the Mayor




(a) PP Mayor

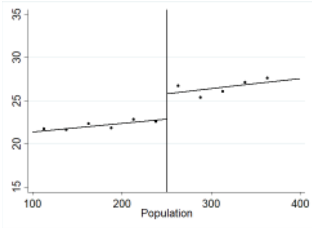



(b) PSOE Mayor

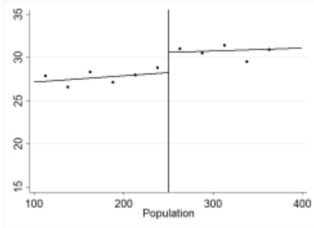
The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 25 p.p.-wide bins of population (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

Figure A6: CL vs OL: The Effect of on the Share of Female Candidates and Councilors, by Party

Share of female candidates (%)

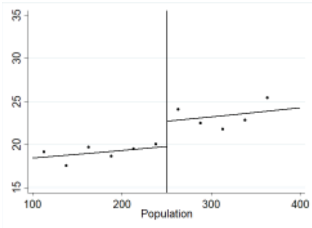



(a) PP

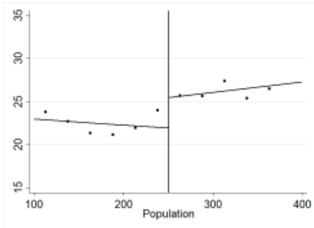



(b) PSOE

Share of female councilors (%)

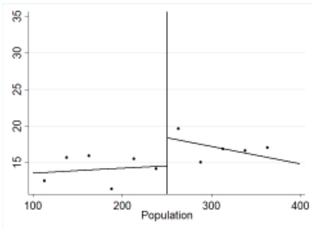



(c) PP

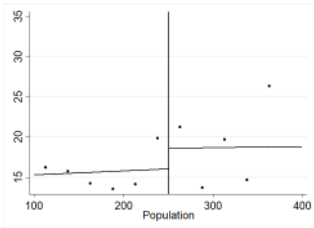



(d) PSOE

Share of female mayors (%)




(e) PP




(f) PSOE

The unit of observation is a municipality-election. Circles represent the local averages of the variables. Averages are calculated within 25 p.p.-wide bins of population (x-axis). Continuous lines are a linear fit over the original (unbinned) data.

# Appendix Tables

Table A1: 1,000-Inhabitant Threshold: Covariate Smoothness around the Threshold

| Variable | (1) Mean | (2) IK Bandwidth | (3) CCT Bandwidth | (4) IK Results | (5) CCT Results |
|---|---|---|---|---|---|
| Average Age (years) | 46.60 | 247 | 147 | -0.108 | -0.184 |
| | | [N = 2903] | [N = 1729] | (0.366) | (0.452) |
| Foreigners (%) | 6.205 | 267 | 84 | 0.236 | -1.113 |
| | | [N = 3132] | [N = 1006] | (0.558) | (0.859) |
| EU Foreigners (%) | 52.13 | 215 | 116 | -1.976 | -3.683 |
| | | [N = 2493] | [N = 1335] | (2.529) | (3.252) |
| Ideology (%) | 2.402 | 317 | 125 | 1.862 | 0.272 |
| | | [N = 3824] | [N = 1467] | (1.522) | (1.776) |
| Turnout (%) | 76.83 | 308 | 136 | 0.288 | 0.503 |
| | | [N = 3708] | [N = 1585] | (0.573) | (0.779) |
| Unemployment (%) | 5.024 | 189 | 117 | 0.427 | 0.671* |
| | | [N = 2194] | [N = 1366] | (0.291) | (0.347) |
| $\mathbb{1}_{year=2003}$ | 0.261 | 323 | 33 | -0.0149 | -0.0186 |
| | | [N = 3921] | [N = 1698] | (0.0271) | (0.0392) |
| $\mathbb{1}_{year=2007}$ | 0.229 | 307 | 37 | 0.0290 | 0.0230 |
| | | [N = 3683] | [N = 1839] | (0.0225) | (0.0349) |
| $\mathbb{1}_{year=2011}$ | 0.259 | 271 | 34 | 0.00529 | -0.00598 |
| | | [N = 3185] | [N = 1518] | (0.0253) | (0.0381) |
| $\mathbb{1}_{year=2015}$ | 0.259 | 277 | 39 | -0.0222 | 0.0192 |
| | | [N = 3276] | [N = 1199] | (0.0282) | (0.0464) |

Column (1) shows the mean of the variables. Columns (2) and (3) show the Imbens and Kalyanaraman (2012) and Calonico, Cattaneo, and Titiunik (2014) bandwidth values and number of observations for a placebo test that estimates the effect of CL on the corresponding variable. Columns (4) and (5) show the results for the placebo tests: each column is a separate local linear regression with a uniform kernel. Standard errors, clustered at both municipality and the running variable, are in parentheses.

Table A2: Effect of CL on the Party of the Mayor

Panel A: PP Mayor

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $CL$ | -0.0113 | -0.0187 | 0.0156 | -0.0297 | -0.030 |
| | (0.0362) | (0.0443) | (0.0307) | (0.0416) | p-v = 0.299 |
| Constant | 0.479 | 0.477 | 0.444 | 0.504 | |
| | (0.0275) | (0.0353) | (0.0226) | (0.0325) | |
| Observations | 4376 | 2135 | 7194 | 7194 | 1342 |
| Bw Size | 98 | 50 | 150 | 150 | 31 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Panel B: PSOE Mayor

| | (1) | (2) | (3) | (4) | (5) |
|---|---|---|---|---|---|
| $CL$ | 0.0153 | 0.0310 | 0.0145 | 0.0175 | 0.024 |
| | (0.0390) | (0.0426) | (0.0285) | (0.0388) | p-v = 0.347 |
| Constant | 0.284 | 0.269 | 0.277 | 0.267 | |
| | (0.0318) | (0.0352) | (0.0213) | (0.0307) | |
| Observations | 2639 | 1729 | 7194 | 7194 | 1342 |
| Bw Size | 61 | 41 | 150 | 150 | 31 |
| Bw Method | IK | CCT | All | All | Local |
| Polynomial | 1 | 1 | 1 | 2 | randomization |

Results from estimating $Outcome_{mt} = \alpha + \beta CL_{mt} + f(Pop_{mt} - 250) + u_{mt}$. Each column reports a separate local polynomial regression estimate with the specified bandwidth and polynomial order. Separate polynomials are fitted on each side of the threshold. The last column shows the results from Cattaneo, Frandsen, and Titiunik (2015)'s procedure. Standard errors clustered by municipality and population in parentheses.

Table A3: Effect of CL on the Share of Female Candidates, Councilors, and Mayors, by Party

| | (1) Candidates | (2) Candidates | (3) Councilors | (4) Councilors | (5) Mayors | (6) Mayors |
|---|---|---|---|---|---|---|
| $CL$ | 3.344** | 3.454** | 3.312* | 3.601* | 6.409 | -0.216 |
| | (1.309) | (1.687) | (1.954) | (1.945) | (3.922) | (5.084) |
| Constant | 22.85 | 27.77 | 21.21 | 21.86 | 15.63 | 20.51 |
| | (1.056) | (1.499) | (1.536) | (1.505) | (2.730) | (3.950) |
| Observations | 5317 | 4975 | 2453 | 6108 | 1541 | 1159 |
| Party | PP | PSOE | PP | PSOE | PP | PSOE |
| p-value: test of equal coefficients | 0.96 | | 0.92 | | 0.31 | |

Results from estimating $Outcome_{mt} = \alpha + \beta CL_{mt} + f(Pop_{mt} - 250) + u_{mt}$. The sample is restricted to PP (columns (1), (3) and (5)) or PSOE (columns (2), (4) and (6)) observations. Each column reports a separate local polynomial regression estimated with the bandwidth from Imbens and Kalyanaraman (2012), being equal to 73 (columns (1) and (2)), 85 (columns (3) and (4)), and 65 (columns (5) and (6)) inhabitants. Separate polynomials are fitted on each side of the threshold. The last column shows the results from Cattaneo, Frandsen, and Titiunik (2015)'s procedure. Standard errors clustered by municipality and population in parentheses.

Table A4: Mayor Selection in OL, by Previous Exposure to a Female Mayor

| | (1) | (2) | (3) |
|---|---|---|---|
| $First$ | 0.179*** | 0.170*** | 0.197** |
| | (0.0391) | (0.0427) | (0.0965) |
| $First*Female1$ | -0.116 | -0.205* | 0.136 |
| | (0.0842) | (0.108) | (0.145) |
| $Female1$ | 0.0749* | 0.127** | -0.0287 |
| | (0.0436) | (0.0567) | (0.0734) |
| $First*Female2$ | 0.218*** | 0.296*** | -0.108 |
| | (0.0741) | (0.0790) | (0.156) |
| $Female2$ | -0.126*** | -0.178*** | 0.0676 |
| | (0.0369) | (0.0389) | (0.0798) |
| Constant | 0.319 | 0.330 | 0.277 |
| | (0.0202) | (0.0223) | (0.0486) |
| Observations | 4430 | 3488 | 942 |
| Sample | All | Not exposed | Exposed |

Results from estimating $y_{cmt} = \theta_0 + \theta_1 \cdot first_{cmt} + g_0(x_{cmt}) + g_1(x_{cmt}) \cdot first_{cmt} + \theta_2 \cdot first_{cmt} \cdot female1_{mt} + \theta_3 female1_{mt} + g_1(x_{cmt}) \cdot first_{cmt} \cdot female1_{mt} + g_1(x_{cmt}) \cdot female1_{mt} + \theta_4 \cdot first_{cmt} \cdot female2_{mt} + \theta_5 female2_{mt} + g_1(x_{cmt}) \cdot first_{cmt} \cdot female2_{mt} + g_1(x_{cmt}) \cdot female2_{mt} + \epsilon_{cmt}$. All the regressions use the IK bandwidth (4.6 p.p.). The sample is restricted to non-incumbent parties. Standard errors clustered by municipality in parentheses.

**BANCO DE ESPAÑA PUBLICATIONS**

WORKING PAPERS

1720 LUIS J. ÁLVAREZ and ANA GÓMEZ-LOSCOS: A menu on output gap estimation methods.

1721 PAULA GIL, FRANCISCO MARTÍ, JAVIER J. PÉREZ, ROBERTO RAMOS and RICHARD MORRIS: The output effects of tax changes: narrative evidence from Spain.

1722 RICARDO GIMENO and ALFREDO IBÁÑEZ: The eurozone (expected) inflation: an option's eyes view.

1723 MIGUEL ANTÓN, SERGIO MAYORDOMO and MARÍA RODRÍGUEZ-MORENO: Dealing with dealers: sovereign CDS comovements.

1724 JOSÉ MANUEL MONTERO: Pricing decisions under financial frictions: evidence from the WDN survey.

1725 MARIO ALLOZA: The impact of taxes on income mobility.

1726 DANILO LEIVA-LEON: Measuring business cycles intra-synchronization in US: a regime-switching interdependence framework.

1727 PIERRE GUÉRIN and DANILO LEIVA-LEON: Model averaging in Markov-Switching models: predicting national recessions with regional data.

1728 MÁXIMO CAMACHO and DANILO LEIVA-LEON: The propagation of industrial business cycles.

1729 JAMES COSTAIN: Costly decisions and sequential bargaining.

1730 MARIO ALLOZA: Is fiscal policy more effective in uncertain times or during recessions?

1731 PIERRE GUÉRIN and DANILO LEIVA-LEON: Monetary policy, stock market and sectoral comovement.

1732 HENRIK JENSEN, IVAN PETRELLA, SØREN HOVE RAVN and EMILIANO SANTORO: Leverage and deepening business cycle skewness.

1733 CÉSAR MARTÍN MACHUCA: External stress early warning indicators.

1734 RODOLFO G. CAMPOS: International migration pressures in the long run.

1735 ANDREA ARIU, ELENA BIEWEN, SVEN BLANK, GUILLAUME GAULIER, MARÍA JESÚS GONZÁLEZ, PHILIPP MEINEN, DANIEL MIRZA, CÉSAR MARTÍN MACHUCA and PATRY TELLO: Firm heterogeneity and aggregate business services exports: micro evidence from Belgium, France, Germany and Spain.

1736 LEONARDO GAMBACORTA, STEFANO SCHIAFFI and ADRIAN VAN RIXTEL: Changing business models in international bank funding.

1737 ENRIQUE MORAL-BENITO and FRANCESCA VIANI: An anatomy of the Spanish current account adjustment: the role of permanent and transitory factors.

1738 MARÍA J. NIETO and LARRY D. WALL: Cross-border banking on the two sides of the Atlantic: does it have an impact on bank crisis management?

1739 JACOPO TIMINI: Currency unions and heterogeneous trade effects: the case of the Latin Monetary Union.

1740 PAULINO FONT, MARIO IZQUIERDO and SERGIO PUENTE: Subsidising mature age employment or throwing coins into a wishing well: a quasi-experimental analysis.

1741 THOMAS FUJIWARA and CARLOS SANZ: Norms in bargaining: evidence from government formation in Spain.

1742 ENRIQUE ALBEROLA, ÁNGEL ESTRADA and FRANCESCA VIANI: Global imbalances from a stock perspective.

1743 ÓSCAR ARCE, RICARDO GIMENO and SERGIO MAYORDOMO: Making room for the needy: the credit-reallocation effects of the ECB's corporate QE.

1744 M. D. GADEA-RIVAS, ANA GÓMEZ-LOSCOS and EDUARDO BANDRÉS: Clustering regional business cycles.

1745 NEZIH GUNER, YULIYA KULIKOVA and JOAN LLULL: Marriage and health: selection, protection, and assortative mating.

1746 SERGIO MAYORDOMO and MARÍA RODRÍGUEZ-MORENO: Did the bank capital relief induced by the supporting factor enhance SME lending?

1747 KATALIN BODNÁR, LUDMILA FADEJEVA, MARCO HOEBERICHTS, MARIO IZQUIERDO PEINADO, CHRISTOPHE JADEAU and ELIANA VIVIANO: Credit shocks and the European Labour market.

1748 PIERRE GUÉRIN, DANILO LEIVA-LEON and MASSIMILIANO MARCELLINO: Markov-switching three-pass regression filter.

1749 ISABEL ARGIMÓN: Decentralized multinational banks and risk taking: the Spanish experience in the crisis.

1750 BING XU: Permissible collateral and access to finance: evidence from a quasi-natural experiment.

1751 GERGELY AKOS GANICS: Optimal density forecast combinations.

1801 OLYMPIA BOVER, LAURA HOSPIDO and ERNESTO VILLANUEVA: The impact of high school financial education on financial knowledge and choices: evidence from a randomized trial in Spain.

1802 IGNACIO HERNANDO, IRENE PABLOS, DANIEL SANTABÁRBARA and JAVIER VALLÉS: Private Saving. New Cross-Country Evidence Based on Bayesian Techniques.

1803 PABLO AGUILAR and JESÚS VÁZQUEZ: Term structure and real-time learning.

1804 MORITZ A. ROTH: International co-movements in recessions.

1805 ANGELA ABBATE and DOMINIK THALER: Monetary policy and the asset risk-taking channel.

1806 PABLO MARTÍN-ACEÑA: Money in Spain. New historical statistics. 1830-1998.

1807 GUILHERME BANDEIRA: Fiscal transfers in a monetary union with sovereign risk.

1808 MIGUEL GARCÍA-POSADA GÓMEZ: Credit constraints, firm investment and growth: evidence from survey data.

1809 LAURA ALFARO, MANUEL GARCÍA-SANTANA and ENRIQUE MORAL-BENITO: On the direct and indirect real effects of credit supply shocks.

1810 ROBERTO RAMOS and CARLOS SANZ: Backing the incumbent in difficult times: the electoral impact of wildfires.

1811 GABRIEL JIMÉNEZ, ENRIQUE MORAL-BENITO and RAQUEL VEGAS: Bank lending standards over the cycle: the role of firms' productivity and credit risk.

1812 JUAN S. MORA-SANGUINETTI and ROK SPRUK: Industry vs services: do enforcement institutions matter for specialization patterns? Disaggregated evidence from Spain.

1813 JAMES CLOYNE, CLODOMIRO FERREIRA and PAOLO SURICO: Monetary policy when households have debt: new evidence on the transmission mechanism.

1814 DMITRI KIRPICHEV and ENRIQUE MORAL-BENITO: The costs of trade protectionism: evidence from Spanish firms and non-tariff measures.

1815 ISABEL ARGIMÓN, CLEMENS BONNER, RICARDO CORREA, PATTY DUIJM, JON FROST, JAKOB DE HAAN, LEO DE HAAN and VIKTORS STEBUNOVS: Financial institutions' business models and the global transmission of monetary policy.

1816 JOSE ASTURIAS, MANUEL GARCÍA-SANTANA and ROBERTO RAMOS: Competition and the welfare gains from transportation infrastructure: evidence from the Golden Quadrilateral of India.

1817 SANDRA GARCÍA-URIBE: Multidimensional media slant: complementarities in news reporting by US newspapers.

1818 PILAR CUADRADO, AITOR LACUESTA, MARÍA DE LOS LLANOS MATEA and F. JAVIER PALENCIA-GONZÁLEZ: Price strategies of independent and branded dealers in retail gas market. The case of a contract reform in Spain.

1819 ALBERTO FUERTES, RICARDO GIMENO and JOSÉ MANUEL MARQUÉS: Extraction of inflation expectations from financial instruments in Latin America.

1820 MARIO ALLOZA, PABLO BURRIEL and JAVIER J. PÉREZ: Fiscal policies in the euro area: revisiting the size of spillovers.

1821 MARTA MARTÍNEZ-MATUTE and ALBERTO URTASUN: Uncertainty, firm heterogeneity and labour adjustments. Evidence from European countries.

1822 GABRIELE FIORENTINI, ALESSANDRO GALESI, GABRIEL PÉREZ-QUIRÓS and ENRIQUE SENTANA: The rise and fall of the natural interest rate.

1823 ALBERTO MARTÍN, ENRIQUE MORAL-BENITO and TOM SCHMITZ: The financial transmission of housing bubbles: evidence from Spain.

1824 DOMINIK THALER: Sovereign default, domestic banks and exclusion from international capital markets.

1825 JORGE E. GALÁN and JAVIER MENCÍA: Empirical assessment of alternative structural methods for identifying cyclical systemic risk in Europe.

1826 ROBERTO BLANCO and NOELIA JIMÉNEZ: Credit allocation along the business cycle: evidence from the latest boom bust credit cycle in Spain.

1827 ISABEL ARGIMÓN: The relevance of currency-denomination for the cross-border effects of monetary policy.

1828 SANDRA GARCÍA-URIBE: The effects of tax changes on economic activity: a narrative approach to frequent anticipations.

1829 MATÍAS CABRERA, GERALD P. DWYER and MARÍA J. NIETO: The G-20 regulatory agenda and bank risk.

1830 JACOPO TIMINI and MARINA CONESA: Chinese exports and non-tariff measures: testing for heterogeneous effects at the product level.

1831 JAVIER ANDRÉS, JOSÉ E. BOSCÁ, JAVIER FERRI and CRISTINA FUENTES-ALBERO: Households' balance sheets and the effect of fiscal policy.

1832 ÓSCAR ARCE, MIGUEL GARCÍA-POSADA, SERGIO MAYORDOMO and STEVEN ONGENA: Adapting lending policies when negative interest rates hit banks' profits.

1833 VICENTE SALAS, LUCIO SAN JUAN and JAVIER VALLÉS: Corporate cost and profit shares in the euro area and the US: the same story?

1834 MARTÍN GONZALEZ-EIRAS and CARLOS SANZ: Women's representation in politics: voter bias, party bias, and electoral systems.

| | |
|---|---|
| BANCO DE **ESPAÑA**<br>Eurosistema | Unidad de Servicios Auxiliares<br>Alcalá, 48 - 28014 Madrid<br>E-mail: publicaciones@bde.es<br>www.bde.es |